

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
10 September 2004 (10.09.2004)

PCT

(10) International Publication Number
WO 2004/075713 A2

- (51) International Patent Classification⁷: **A61B**
- (21) International Application Number:
PCT/CA2004/000281
- (22) International Filing Date: 26 February 2004 (26.02.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/450,406 26 February 2003 (26.02.2003) US
- (71) Applicant (for all designated States except US): **MOUNT SINAI HOSPITAL** [CA/CA]; 600 University Avenue, Room 970, Toronto, Ontario M5G 1X5 (CA).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **DIAMANDIS, Eleftherios, P.** [CA/CA]; 44 Gerrard Street West, Suite 1504, Toronto, Ontario M5G 2K2 (CA).
- (74) Agent: **KURDYDYK, Linda**; McCarthy Tetrault, 66 Wellington Street West, Suite 4700, P.O. Box 48, Toronto Dominion Bank Tower (CA).

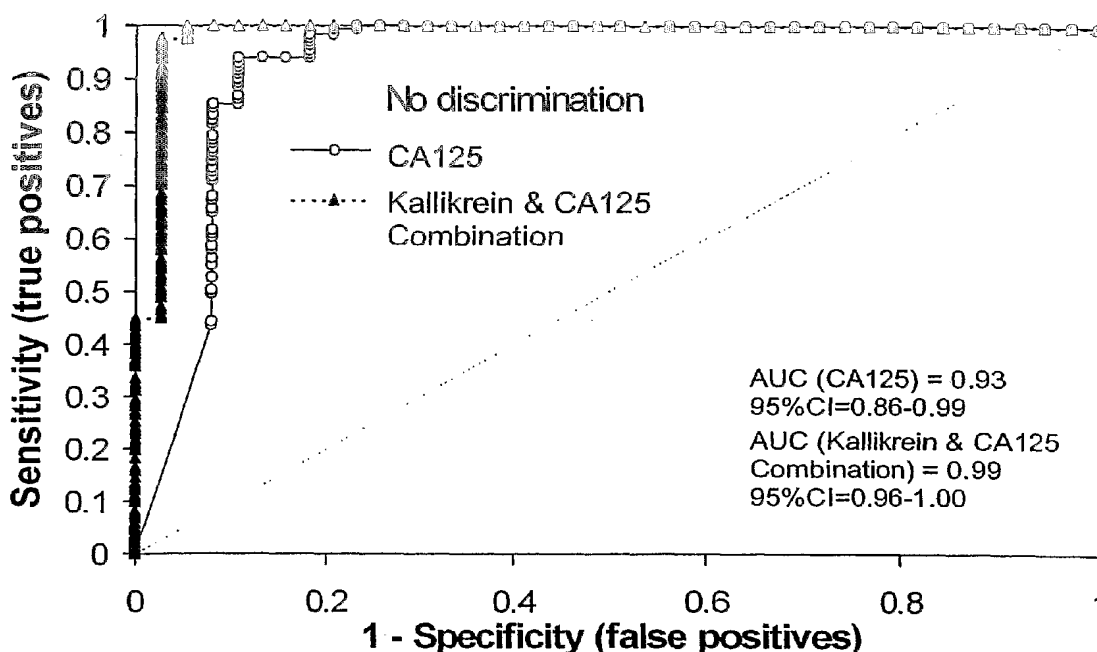
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: MULTIPLE MARKER ASSAY FOR DETECTION OF OVARIAN CANCER



(57) Abstract: Methods for diagnosing and monitoring ovarian cancer in a subject comprising measuring a plurality of kallikrein polypeptides, and optionally CA125, or nucleic acids encoding the polypeptides in a sample from the subject. The kallikrein polypeptides include kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

WO 2004/075713 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

TITLE: Multiple Marker Assay for Detection of Ovarian Cancer**FIELD OF THE INVENTION**

The invention relates to compositions, kits, and methods for detecting, characterizing, preventing, and treating ovarian cancer.

5 **BACKGROUND OF THE INVENTION**

Epithelial ovarian carcinoma is the most common and most lethal of all gynecologic malignancies. Only 30% of ovarian tumors are diagnosed at an early stage (Stage I/II), when survival rates reach 90%. The rest are diagnosed at an advanced stage, with survival rates of less than 20% (Greenlee RT, Hill-Harmon MB, Murray T, et al., 2001. *CA Cancer J Clin* .2001;51:15-36). Currently, the only well-accepted
10 serological marker is CA125, a large glycoprotein of unknown function (Meyer T, Rustin GJ., *Br J Cancer* .2000;82:1535-1538). However, CA125 has limitations as a diagnostic, prognostic and screening tool (Holschneider CH, Berek JS, *Semin Surg Oncol* .2000;19:3-10). Consequently, there is a need to enhance the overall diagnostic/prognostic capability of CA125.

Kallikreins are a subgroup of secreted serine proteases, encoded by highly conserved and tightly
15 clustered multigene families in humans, rats and mice. The human kallikrein gene family resides on chromosome 19q13.4 and is comprised of 15 members, whose genes are designated as *KLK1* to *KLK15* and the corresponding proteins as hK1 to hK15 (Yousef GM, Diamandis EP., *Endocr Rev* .2001;22:184-204l; Yousef GM, Chang A, Scorilas A, et al., *Biochem Biophys Res Commun*. 2000;276:125-133; Diamandis EP, Yousef GM, Clements J, et al. *Clin Chem* .2000;46:1855-1858). Kallikreins are expressed in a wide variety
20 of tissues and are found in many biological fluids (e.g. cerebrospinal fluid, serum, seminal plasma, milk, etc.) where they are predicted to process specific substrates. Kallikreins may participate in cascade reactions similar to those involved in digestion, fibrinolysis, coagulation, wound healing and apoptosis ((Yousef GM, Diamandis EP., *Endocr Rev* .2001;22:184-204l). Many kallikreins have been found to be differentially expressed in endocrine-related malignancies (Diamandis EP, Yousef GM, *Expert Rev. Mol. Diagn*
25 .2001;1:182-190), including prostate (Barry MJ. Clinical practice, *N Engl J Med* .2001;344:1373-1377; Rittenhouse HG, Finlay JA, Mikolajczyk SD, et al., *Crit Rev Clin Lab Sci* .1998;35:275-368; and Yousef GM, Scorilas A, Jung K, et al., *J Biol Chem* .2001;276:53-61), ovarian (Kim H, Scorilas A, Katsaros D, et al., *Br J Cancer*, 2001;84:643-650; Anisowicz A, Sotiropoulou G, Stenman, et al., *Mol Med* .1996;2:624-636; Tanimoto H, Underwood LJ, Shigemasa K, et al., *Cancer* .1999;86:2074-2082; Magklara A, Scorilas A, Katsaros D, et al., *Clin Cancer Res* .2001;7:806-811; Yousef GM, Kyriakopoulou LG, Scorilas A, et al.,
30 *Cancer Res* .2001;61:7811-7818; Luo L, Bunting P, Scorilas A, Diamandis EP., *Clin Chim Acta* .2001;306:111-118), breast (Yousef GM, Magklara A, Chang A, et al., *Cancer Res* .2001;61:3425-3431; Yousef GM, Chang A, Diamandis EP; *J Biol Chem* .2000;275:11891-11898; and Yousef GM, Magklara A, Diamandis EP, *Genomics* .2000;69:331-341), and testicular cancer (Luo LY, Rajpert-De Meyts ER, Jung K, et al.,2001;85:220-224). In addition, many kallikrein genes examined thus far are under steroid hormone regulation, implicating a role for kallikreins in endocrine-related tissues (Yousef GM, Diamandis EP., *Endocr Rev*., 2001;22:184-204). Furthermore, hK6, hK10 and hK11 have been recently identified as novel serological ovarian cancer biomarkers (Luo L, Bunting P, Scorilas A, Diamandis EP., *Clin Chim Acta* .2001;306:111-118 Diamandis EP, Yousef GM, Soosaipillai AR, Bunting P., *Clin Biochem*. 2000;33:579-

- 2 -

583, and Diamandis EP, Okui A, Mitsui S, et al., *Cancer Res* .2002;62:295-300).

SUMMARY OF THE INVENTION

The present invention seeks to overcome the drawbacks inherent in the prior art and seeks to provide sensitive and accurate multimarker methods for the detection of ovarian cancer. A plurality of kallikrein polypeptides and polynucleotides encoding the polypeptides, optionally in combination with CA125 and polynucleotides encoding CA125 can have particular application in the detection of ovarian cancer. A plurality of kallikrein markers (i.e. two or more of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11) and polynucleotides encoding the polypeptides, optionally in combination with CA125 and polynucleotides encoding CA125, constitute biomarkers for the diagnosis, monitoring, progression, treatment, and prognosis of ovarian cancer, and they may be used as biomarkers before surgery or after relapse.

In accordance with the methods of the invention, the presence of levels of markers in a sample can be assessed, for example by detecting the presence in the sample of (a) polypeptides or polypeptide fragments corresponding to the markers; (b) metabolites which are produced directly or indirectly by polypeptides corresponding to the markers; (c) transcribed nucleic acids or fragments thereof having at least a portion with which the markers are substantially identical; and/or (c) transcribed nucleic acids or fragments thereof, wherein the nucleic acids hybridize with the markers.

In an aspect of the invention, a method is provided for detecting ovarian cancer in a patient comprising detecting a plurality of kallikrein polypeptides, optionally in combination with CA125, in a sample from the patient wherein the method provides substantially increased sensitivity compared to methods using CA125 alone. In an embodiment, sensitivity is increased by at least 0.5%, 1%, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 25%, 30%, and 35% compared to using CA125 alone.

In an embodiment, the invention provides a method for detecting a plurality of kallikrein markers, and optionally CA125, associated with ovarian cancer in a patient comprising:

- (a) obtaining a sample from a patient;
- (b) detecting or identifying in the sample kallikrein markers, optionally in combination with CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- (c) comparing the detected amounts with amounts detected for a standard.

The term "detect" or "detecting" includes assaying, assessing, imaging or otherwise establishing the presence or absence of the target kallikrein and CA125 polypeptides or polynucleotides encoding the polypeptides, subunits thereof, or combinations of reagent bound targets, and the like, or assaying for, imaging, ascertaining, establishing, or otherwise determining one or more factual characteristics of ovarian cancer, metastasis, stage, or similar conditions. The term encompasses diagnostic, prognostic, and monitoring applications. The kallikrein polypeptides and CA125 can be detected individually, sequentially, or simultaneously.

According to a method involving kallikrein markers optionally in combination with CA125, the levels in the sample of the kallikrein markers (2, 3, 4, 5, or 6) and optionally CA125, wherein the markers comprise or are selected from kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein

- 3 -

11, are compared with the normal levels of the kallikrein markers, and optionally CA125, in samples of the same type obtained from controls (e.g. samples from individuals not afflicted with ovarian cancer). Significantly different levels in the sample of the kallikrein markers (and optionally CA125) relative to the normal levels in a control is indicative of ovarian cancer.

5 In an embodiment, the invention provides a method for diagnosing and monitoring ovarian carcinoma in a subject comprising detecting in a sample from the subject kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. The kallikrein markers and CA125 can be detected using antibodies that bind to the kallikrein markers and CA125 or parts thereof.

10 Thus, the invention provides a method of assessing whether a patient is afflicted with or has a pre-disposition for ovarian cancer, the method comprising comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a sample from the patient, wherein the kallikrein markers comprise kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 15 (b) normal levels of kallikrein markers, and optionally CA125, in samples of the same type obtained from control patients not afflicted with ovarian cancer, wherein significantly different levels of the kallikrein markers and optionally CA125, relative to the corresponding normal levels of the kallikrein markers, and optionally CA125, is an indication that the patient is afflicted with ovarian cancer.

20 In an embodiment of a method of assessing whether a patient is afflicted with ovarian cancer (e.g. screening, detection of a recurrence, reflex testing), the method comprises comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 25 (b) normal levels of the kallikrein markers, and optionally CA125, in a control non-ovarian cancer sample.

A significant difference between the levels of the kallikrein markers, and optionally CA125, in the patient sample and the normal levels is an indication that the patient is afflicted with ovarian cancer.

30 The invention further relates to a method of assessing the efficacy of a therapy for inhibiting ovarian cancer in a patient. This method comprises comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a first sample obtained from the patient prior to providing at least a portion of the therapy to the patient, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11 ; and
- 35 (b) levels of the kallikrein markers, and optionally CA125, in a second sample obtained from the patient following therapy.

A significant difference between the levels of the kallikrein markers, and optionally CA125, in the second sample, relative to the first sample, is an indication that the therapy is efficacious for inhibiting ovarian cancer.

- 4 -

The "therapy" may be any therapy for treating ovarian cancer including but not limited to chemotherapy, immunotherapy, gene therapy, radiation therapy, and surgical removal of tissue. Therefore, the method can be used to evaluate a patient before, during, and after therapy, for example, to evaluate the reduction in tumor burden.

5 In an aspect, the invention provides a method for monitoring the progression of ovarian cancer in a patient, the method comprising:

- (a) detecting in a patient sample at a first time point, kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 10 (b) repeating step (a) at a subsequent point in time; and
- (c) comparing the levels detected in (a) and (b), and therefrom monitoring the progression of ovarian cancer in the patient.

In another aspect, the invention provides a method for assessing the aggressiveness or indolence of ovarian cancer (e.g. staging), the method comprising comparing:

- 15 (a) levels of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- (b) normal levels of the kallikrein markers, and optionally CA125 in a control sample.

A significant difference between the levels in the sample and the normal levels is an indication that
20 the cancer is aggressive or indolent.

The invention provides a method for determining whether ovarian cancer has metastasized or is likely to metastasize in the future, the method comprising comparing:

- (a) levels of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
- 25 (b) normal levels (or non-metastatic levels) of the kallikrein markers, and optionally CA125, in a control sample.

A significant difference between the levels in the patient sample and the normal levels is an indication that the cancer has metastasized or is likely to metastasize in the future.

30 The invention also provides a method for assessing the potential efficacy of a test agent for inhibiting ovarian cancer in a patient, and a method of selecting an agent for inhibiting ovarian cancer in a patient.

The invention further provides a method of inhibiting ovarian cancer in a patient comprising:

- (a) obtaining a sample comprising cancer cells from the patient;
- 35 (b) separately maintaining aliquots of the sample in the presence of a plurality of test agents;
- (c) comparing levels of kallikrein markers, and optionally CA125, in each of the aliquots, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11;
- (d) administering to the patient at least one of the test agents which alters the levels of the

- 5 -

kallikrein markers, and optionally CA125, in the aliquot containing that test agent, relative to other test agents.

The invention also contemplates a method of assessing the ovarian carcinogenic potential of a test compound comprising:

- 5 (a) maintaining separate aliquots of ovarian cells in the presence and absence of the test compound; and
- (b) comparing levels of kallikrein markers, and optionally CA125, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

10 A significant difference between the levels of the kallikrein markers, and optionally CA125, in the aliquot maintained in the presence of (or exposed to) the test compound relative to the aliquot maintained in the absence of the test compound, indicates that the test compound possesses ovarian carcinogenic potential.

In preferred embodiments of the methods of the invention, the kallikrein markers comprise a plurality of kallikrein markers, for example, at least three, four, five, or six of the markers. In particular, a
15 plurality of kallikrein markers may be selected from the group consisting of kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10, from the group consisting of kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, or from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

Other methods of the invention employ one or more polynucleotides capable of hybridizing to
20 polynucleotides encoding kallikrein markers, and optionally CA125. Methods for detecting polynucleotides encoding a kallikrein markers, and optionally CA125, can be used to monitor ovarian cancer by detecting the nucleic acids.

Thus, the present invention relates to a method for diagnosing and monitoring ovarian cancer in a sample from a subject comprising isolating nucleic acids, preferably mRNA, from the sample; and detecting
25 polynucleotides encoding kallikrein markers, and optionally CA125, in the sample. The presence of different levels of polynucleotides encoding kallikrein markers, and optionally CA125, in the sample compared to a standard or control is indicative of disease, disease stage, and/or prognosis, e.g. longer progression-free and overall survival.

In an embodiment, the invention provides methods for determining the presence or absence of
30 ovarian cancer in a subject comprising (a) contacting a sample obtained from the subject with oligonucleotides that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125; and (b) detecting in the sample levels of nucleic acids that hybridize to the polynucleotides relative to a predetermined cut-off value, and therefrom determining the presence or absence of ovarian cancer in the subject. Within certain embodiments, mRNA is detected via polymerase chain reaction using, for example
35 oligonucleotide primers that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, or complements of such polynucleotides. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing oligonucleotide probes that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, or complements of such polynucleotides.

When using mRNA detection, the method may be carried out by combining isolated mRNA with

- 6 -

reagents to convert to cDNA according to standard methods; treating the converted cDNA with amplification reaction reagents (such as cDNA PCR reaction reagents) in a container along with an appropriate mixture of nucleic acid primers; reacting the contents of the container to produce amplification products; and analyzing the amplification products to detect the presence of polynucleotides encoding kallikrein markers, and optionally CA125, in the sample. For mRNA the analyzing step may be accomplished using Northern Blot analysis to detect the presence of polynucleotides encoding kallikrein markers, and optionally CA125. The analysis step may be further accomplished by quantitatively detecting the presence of polynucleotides encoding kallikrein markers, and optionally CA125, in the amplification product, and comparing the quantity of markers detected against a panel of expected values for the known presence or absence of the kallikrein markers in normal and malignant tissue derived using similar primers.

In embodiments of the methods of the invention, a plurality (eg. three, four, five or six) polynucleotides encoding kallikrein polypeptides are employed. In particular, a plurality of polynucleotides encoding kallikrein markers may be selected from the group consisting of polynucleotides encoding (i) kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10; (ii) polynucleotides encoding kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (iii) polynucleotides encoding kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

The invention also provides a diagnostic composition comprising a plurality of kallikrein polypeptides and optionally CA125 polypeptide, or polynucleotides encoding the polypeptides, or agents that bind to the polypeptides or polynucleotides.

In an embodiment, the composition comprises probes that specifically hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, or fragments thereof. In another embodiment a composition is provided comprising specific primer pairs capable of amplifying polynucleotides encoding kallikrein markers, and optionally CA125, using polymerase chain reaction methodologies. In a still further embodiment, the composition comprises agents that bind to kallikrein markers, and optionally CA125, (e.g. antibodies) or fragments thereof. Probes, primers, and agents can be labeled with detectable substances.

In an aspect the invention provides an *in vivo* method comprising administering to a subject agents that have been constructed to target kallikrein markers, and optionally CA125.

The invention therefore contemplates an *in vivo* method comprising administering to a mammal imaging agents that carry labels for imaging and that bind to kallikrein markers, and optionally CA125, and then imaging the mammal.

Still further the invention relates to therapeutic applications for ovarian cancer employing kallikrein markers, and optionally CA125, nucleic acids encoding the polypeptides, and/or agents identified using methods of the invention.

The invention also includes kits for carrying out methods of the invention. In an embodiment, the kit is for assessing whether a patient is afflicted with ovarian cancer and it comprises reagents for assessing kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

In another aspect the invention relates to a kit for assessing the suitability of each of a plurality of test compounds for inhibiting ovarian cancer in a patient. The kit comprises reagents for assessing kallikrein

- 7 -

markers, and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. The kit may also comprise a plurality of test agents or compounds.

5 The invention contemplates a kit for assessing the presence of ovarian cancer cells, wherein the kit comprises antibodies specific for selected kallikrein markers, and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

10 Additionally the invention provides a kit for assessing the ovarian carcinogenic potential of a test compound. The kit comprises ovarian cells and reagents for assessing kallikrein markers, and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

15 In an aspect the invention provides a method of treating a patient afflicted with ovarian cancer comprising providing to cells of a patient antisense oligonucleotides complementary to polynucleotides encoding kallikrein markers, and optionally CA125, which are overexpressed in ovarian cancer. In an alternative method, expression of genes corresponding to kallikrein markers, and optionally CA125, which are underexpressed in ovarian cancer are increased.

20 The invention relates to a method of inhibiting ovarian cancer in a patient at risk for developing ovarian cancer comprising inhibiting or increasing expression (or overexpression) of genes encoding kallikrein markers and optionally CA125, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, that are either overexpressed or underexpressed, in ovarian cancer.

25 Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples while indicating preferred embodiments of the invention are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

DESCRIPTION OF THE DRAWINGS

The invention will now be described in relation to the drawings in which

30 Figure 1 is a graph showing hk5 concentration in serum from non-cancer and cancer patients.

Figure 2 is a graph showing hk6 concentration in serum from non-cancer and cancer patients.

Figure 3 is a graph showing hk7 concentration in serum from non-cancer and cancer patients.

Figure 4 is a graph showing hk8 concentration in serum from non-cancer and cancer patients.

Figure 5 is a graph showing hk10 concentration in serum from non-cancer and cancer patients.

Figure 6 is a graph showing hk11 concentration in serum from non-cancer and cancer patients.

35 Figure 7 is a graph showing CA125 concentration in serum from non-cancer and cancer patients.

Figure 8 is a ROC curve illustrating the added value of using kallikreins and CA125 together in a multivariate function.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to newly discovered correlations between expression of certain markers and

ovarian cancer. The combinations of markers described herein may provide sensitive methods for detecting ovarian cancer. The levels of expression of a combination of markers described herein may correlate with the presence of ovarian cancer or a pre-malignant condition in a patient. Methods are provided for detecting the presence of ovarian cancer in a sample, the absence of ovarian cancer in a sample, the stage of an ovarian cancer, the grade of an ovarian cancer, the benign or malignant nature of an ovarian cancer, the metastatic potential of an ovarian cancer, assessing the histological type of neoplasm associated with the ovarian cancer, the indolence or aggressiveness of the cancer, and other characteristics of ovarian cancer that are relevant to prevention, diagnosis, characterization, and therapy of ovarian cancer in a patient. Methods are also provided for assessing the efficacy of one or more test agents for inhibiting ovarian cancer, assessing the efficacy of a therapy for ovarian cancer, monitoring the progression of ovarian cancer, selecting an agent or therapy for inhibiting ovarian cancer, treating a patient afflicted with ovarian cancer, inhibiting ovarian cancer in a patient, and assessing the carcinogenic potential of a test compound.

Glossary

The terms "sample", "biological sample", and the like, mean a material known or suspected of expressing or containing a plurality of kallikrein markers or polypeptides (2, 3, 4, 5, or 6 polypeptides), and optionally CA125 polypeptide, or polynucleotides encoding the polypeptides. The test sample can be used directly as obtained from the source or following a pretreatment to modify the character of the sample. The sample can be derived from any biological source, such as tissues, extracts, or cell cultures, including cells (e.g. tumor cells), cell lysates, and physiological fluids, such as, for example, whole blood, plasma, serum, saliva, ocular lens fluid, cerebral spinal fluid, sweat, urine, milk, ascites fluid, synovial fluid, peritoneal fluid and the like. The sample can be obtained from animals, preferably mammals, most preferably humans. The sample can be treated prior to use, such as preparing plasma from blood, diluting viscous fluids, and the like. Methods of treatment can involve filtration, distillation, extraction, concentration, inactivation of interfering components, the addition of reagents, and the like. Nucleic acids and polypeptides may be isolated from the samples and utilized in the methods of the invention. In a preferred embodiment, the sample is a serum sample.

The term "subject" or "patient" refers to a warm-blooded animal such as a mammal, which is suspected of having ovarian cancer, or a condition, disease, or syndrome associated with ovarian cancer. Preferably, "subject" refers to a human.

"CA125", "CA125 polypeptide", or "carbohydrate antigen 125" refers to a high-molecular weight mucin, which can be defined by its ability to bind to monoclonal antibody OC125. The CA125 protein core comprises a short cytoplasmic core tail, a transmembrane domain, and a large and heavily glycosylated extracellular domain dominated by a repeat domain of 156 amino acids rich in serine, threonine, and proline (Yin BW and Lloyd KO, J Biol Chem. 2001, 276:27371-27375; O'Brian TJ et al, Tumor Biol., 2001 22:348-366; and Hovig E. et al, Tumor Biol. 2001, 22:345-347). The sequence of CA125 is shown in GenBank Accession No. NP_078966, AAL65133 and AF414442 (SEQ ID NO. 1). The term includes the native-sequence polypeptides, isoforms, precursors and chimeric polypeptides. The term also includes the native sequence polypeptide, including polypeptide variants and polypeptides with substantial sequence identity (e.g. at least about 45%, preferably 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or

99% sequence identity) to the sequence of GenBank Accession No.NP_078966 (SEQ ID NO. 1), and that preferably retain the immunogenic activity of the corresponding native sequence polypeptide.

“Kallikrein polypeptides” or “kallikrein markers” comprise kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. The term includes the native-sequence polypeptides, isoforms, precursors and chimeric polypeptides. The amino acid sequences for native kallikrein polypeptides employed in the present invention include the sequences found in GenBank for each polypeptide as shown in Table 1, and in SEQ ID NO: 3 (kallikrein 5), NO.6 (kallikrein 6), NO. 10 (kallikrein 7), NO. 13 (kallikrein 8), NO. 16 (kallikrein 10), and NOs. 19 and 20 (kallikrein 11), or a portion thereof. Other useful polypeptides are substantially identical to these sequences (e.g. at least about 45%, preferably 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity), and preferably retain the immunogenic activity of the corresponding native-sequence kallikrein polypeptide.

A "native-sequence polypeptide" comprises a polypeptide having the same amino acid sequence of a polypeptide derived from nature. Such native-sequence polypeptides can be isolated from nature or can be produced by recombinant or synthetic means.

The term "native-sequence polypeptide" specifically encompasses naturally occurring truncated or secreted forms of a polypeptide, polypeptide variants including naturally occurring variant forms (e.g., alternatively spliced forms or splice variants), and naturally occurring allelic variants.

The term "polypeptide variant" means a polypeptide having at least about 70-80%, preferably at least about 85%, more preferably at least about 90%, most preferably at least about 95% amino acid sequence identity with a native-sequence polypeptide, in particular having at least 70-80%, 85%, 90%, 95% amino acid sequence identity to the sequences identified in the GenBank Accession Nos. in Table 1 and Accession No. NP_078966, AF414442 and AAL65133 and shown in SEQ ID NOS: 1, 2, 3, 6, 10, 13, 16, 19 and 20. Such variants include, for instance, polypeptides wherein one or more amino acid residues are added to, or deleted from, the N- or C-terminus of the full-length or mature sequences of SEQ ID NOS: 1, 2, 3, 6, 10, 13, 16, 19 and 20, including variants from other species, but excludes a native-sequence polypeptide.

An allelic variant may also be created by introducing substitutions, additions, or deletions into a nucleic acid encoding a native polypeptide sequence such that one or more amino acid substitutions, additions, or deletions are introduced into the encoded protein. Mutations may be introduced by standard methods, such as site-directed mutagenesis and PCR-mediated mutagenesis. In an embodiment, conservative substitutions are made at one or more predicted non-essential amino acid residues. A “conservative amino acid substitution” is one in which an amino acid residue is replaced with an amino acid residue with a similar side chain. Amino acids with similar side chains are known in the art and include amino acids with basic side chains (e.g. Lys, Arg, His), acidic side chains (e.g. Asp, Glu), uncharged polar side chains (e.g. Gly, Asp, Glu, Ser, Thr, Tyr and Cys), nonpolar side chains (e.g. Ala, Val, Leu, Iso, Pro, Trp), beta-branched side chains (e.g. Thr, Val, Iso), and aromatic side chains (e.g. Tyr, Phe, Trp, His). Mutations can also be introduced randomly along part or all of the native sequence, for example, by saturation mutagenesis. Following mutagenesis the variant polypeptide can be recombinantly expressed and the activity of the polypeptide may be determined.

Polypeptide variants include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of a native polypeptide which include fewer amino acids than the full length polypeptides. A portion of a polypeptide can be a polypeptide which is for example, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more amino acids in length. Portions in which regions of a polypeptide are deleted can be prepared by recombinant techniques and can be evaluated for one or more functional activities such as the ability to form antibodies specific for a polypeptide.

A naturally occurring allelic variant may contain conservative amino acid substitutions from the native polypeptide sequence or it may contain a substitution of an amino acid from a corresponding position in a CA125 or kallikrein polypeptide homolog, for example, the murine CA125 or kallikrein polypeptide.

Percent identity of two amino acid sequences, or of two nucleic acid sequences identified herein is defined as the percentage of amino acid residues or nucleotides in a candidate sequence that are identical with the amino acid residues in a CA125 or kallikrein polypeptide or nucleic acid sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid or nucleic acid sequence identity can be achieved in various conventional ways, for instance, using publicly available computer software including the GCG program package (Devereux J. et al., Nucleic Acids Research 12(1): 387, 1984); BLASTP, BLASTN, and FASTA (Atschul, S.F. et al. J. Molec. Biol. 215: 403-410, 1990). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S. et al. NCBI NLM NIH Bethesda, Md. 20894; Altschul, S. et al. J. Mol. Biol. 215: 403-410, 1990). Skilled artisans can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Methods to determine identity and similarity are codified in publicly available computer programs.

CA125 and kallikrein polypeptides include chimeric or fusion proteins. A "chimeric protein" or "fusion protein" comprises all or part (preferably biologically active) of a CA125 or kallikrein polypeptide operably linked to a heterologous polypeptide (i.e., a polypeptide other than the same CA125 or kallikrein polypeptide). Within the fusion protein, the term "operably linked" is intended to indicate that the CA125 or kallikrein polypeptide and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the N-terminus or C-terminus of the CA125 or kallikrein polypeptide. A useful fusion protein is a GST fusion protein in which a kallikrein polypeptide is fused to the C-terminus of GST sequences. Another example of a fusion protein is an immunoglobulin fusion protein in which all or part of a CA125 or kallikrein polypeptide is fused to sequences derived from a member of the immunoglobulin protein family. Chimeric and fusion proteins can be produced by standard recombinant DNA techniques.

CA125 and kallikrein polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods, or by any combination of these and similar techniques.

"CA125 polynucleotides" or "polynucleotides encoding CA125" include nucleic acids that encode a native-sequence polypeptide, a polypeptide variant including a portion of a CA125 polypeptide, an isoform, precursor, and chimeric polypeptide. A nucleic acid sequence encoding native CA125 employed in the

- 11 -

present invention includes the nucleic acid sequence in GenBank Accession No. AF414442 and SEQ ID NO. 2, or a fragment thereof.

“Kallikrein polynucleotides” or “polynucleotides encoding kallikrein markers/polypeptides” refers to kallikrein 5 nucleic acids (KLK5), kallikrein 6 nucleic acids (KLK6), kallikrein 7 nucleic acids (KLK7),
5 kallikrein 8 nucleic acids (KLK8), kallikrein 10 nucleic acids (KLK10), and/or kallikrein 11 nucleic acids (KLK11). The term includes nucleic acids that encode a native-sequence polypeptide, a polypeptide variant including a portion of a kallikrein polypeptide, an isoform, precursor, and chimeric polypeptide.

The polynucleotide sequences encoding native kallikrein polypeptides employed in the present invention include the nucleic acid sequences of the GenBank Accession Nos. identified in Table 1, and in
10 SEQ ID NOS: 4 and 5 (KLK5), NOS. 7, 8, and 9 (KLK6), NOS. 11 and 12 (KLK 7), NOS. 14 and 15 (KLK8), NOS. 17 and 18 (KLK10), and NOS. 21 and 22 (KLK11), or a fragment thereof.

Polynucleotides encoding kallikrein polypeptides and CA125 include nucleic acid sequences complementary to these polynucleotides, and polynucleotides that are substantially identical to these sequences (e.g. at least about 45%, preferably 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%,
15 98%, or 99% sequence identity).

CA125 and kallikrein polynucleotides also include sequences which differ from a nucleic acid sequence of GenBank Accession Nos. identified in Table 1 and SEQ ID NOS: 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 17, 18, 21, and 22, due to degeneracy in the genetic code. As one example, DNA sequence polymorphisms within the nucleotide sequence of a CA125 or kallikrein polypeptide may result in silent mutations which do
20 not affect the amino acid sequence. Variations in one or more nucleotides may exist among individuals within a population due to natural allelic variation. DNA sequence polymorphisms may also occur which lead to changes in the amino acid sequence of CA125 or a kallikrein polypeptide.

CA125 and kallikrein polynucleotides also include nucleic acids that hybridize under stringent conditions, preferably high stringency conditions to a nucleic acid sequence of the GenBank Accession Nos. identified in Table 1 and SEQ ID NOS: 2, 4, 5, 7, 8, 9, 11, 12, 14, 15, 17, 18, 21, and 22. Appropriate
25 stringency conditions which promote DNA hybridization are known to those skilled in the art, or can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C may be employed. The stringency may be selected based on the conditions used in the wash step. By
30 way of example, the salt concentration in the wash step can be selected from a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be at high stringency conditions, at about 65°C.

CA125 and kallikrein polynucleotides also include truncated nucleic acids or fragments and variant forms of the polynucleotides that arise by alternative splicing of an mRNA corresponding to a DNA.

35 The CA125 and kallikrein polynucleotides are intended to include DNA and RNA (e.g. mRNA) and can be either double stranded or single stranded. A polynucleotide may, but need not, include additional coding or non-coding sequences, or it may, but need not, be linked to other molecules and/or carrier or support materials. The polynucleotides for use in the methods of the invention may be of any length suitable for a particular method.

A plurality of kallikrein polypeptides or kallikrein polynucleotides are generally detected in the present invention. "Plurality" refers to 2, 3, 4, 5, or 6 kallikrein polypeptides or polynucleotides, in particular 3, 4, 5, or 6, preferably 4, 5, or 6, more preferably 5 or 6 kallikrein polypeptides or polynucleotides.

In an embodiment a plurality of kallikrein polypeptides is selected from the group consisting of
 5 kallikrein 5, kallikrein 7, and kallikrein 8; kallikrein 5, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; or kallikrein 5, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11. In another embodiment, a plurality of kallikrein polypeptides is selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.

10 In an embodiment, a plurality of kallikrein polynucleotides is selected from the group consisting of KLK5, KLK7, and KLK8; KLK5, KLK8 and KLK10; KLK7, KLK8 and KLK10; KLK5, KLK7, KLK8, and KLK10; KLK7, KLK8, KLK10 and KLK11, or KLK5, KLK7, KLK8, KLK10 and KLK11. In another embodiment, a plurality of kallikrein polynucleotides is selected from the group consisting of KLK5, KLK6, KLK7, KLK8, KLK10, and KLK11.

15 **General Methods**

A variety of methods can be employed for the diagnostic and prognostic evaluation of ovarian cancer involving kallikrein polypeptides, and optionally CA125 polypeptide, and polynucleotides encoding the polypeptides, and the identification of subjects with a predisposition to such disorders. Such methods may, for example, utilize polynucleotides encoding kallikrein polypeptides, and optionally CA125, and
 20 fragments thereof, and binding agents (e.g. antibodies aptamers) against kallikrein polypeptides, and optionally CA125 polypeptide, including peptide fragments. In particular, the polynucleotides and antibodies may be used, for example, for (1) the detection of either over- or under-expression of kallikrein polynucleotides, and optionally CA125, relative to a non-disorder state; and (2) the detection of either an over- or an under-abundance of kallikrein polypeptides, and optionally CA125, relative to a non-disorder
 25 state or the presence of modified (e.g., less than full length) kallikrein polypeptides, and optionally CA125, that correlate with a disorder state, or a progression toward a disorder state.

The invention also contemplates a method for detecting ovarian cancer comprising producing a profile of levels of a plurality of kallikrein markers, and optionally CA125, in cells from a patient, wherein the markers are kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and
 30 comparing the profile with a reference to identify a protein profile for the test cells indicative of disease.

The methods described herein may be used to evaluate the probability of the presence of malignant or pre-malignant cells, for example, in a group of cells freshly removed from a host. Such methods can be used to detect tumors, quantitate their growth, and help in the diagnosis and prognosis of disease. The methods can be used to detect the presence of cancer metastasis, as well as confirm the absence or removal
 35 of all tumor tissue following surgery, cancer chemotherapy, and/or radiation therapy. They can further be used to monitor cancer chemotherapy and tumor reappearance.

The methods described herein can be adapted for diagnosing and monitoring ovarian cancer by detecting a plurality of kallikrein polypeptides, and optionally CA125 polypeptide, or nucleic acids encoding the polypeptides in biological samples from a subject. These applications require that the amount of

- polypeptides or nucleic acids quantitated in a sample from a subject being tested be compared to a predetermined standard. The standard may correspond to levels quantitated for another sample or an earlier sample from the subject, or levels quantitated for a control sample. Levels for control samples from healthy subjects or ovarian cancer subjects may be established by prospective and/or retrospective statistical studies.
- 5 Healthy or normal subjects who have no clinically evident disease or abnormalities may be selected for statistical studies. Diagnosis may be made by a finding of statistically different levels of a plurality of kallikrein polypeptides, and optionally CA125, or nucleic acids encoding same, compared to a control sample or previous levels quantitated for the same subject. A "significant difference" in levels of kallikrein markers or polynucleotides encoding the kallikrein markers in a patient sample compared to a control or
- 10 standard (e.g. normal levels or levels in other samples from a patient) may represent levels that are higher or lower than the standard error of the detection assay, preferably the levels are at least about 1.5, 2, 3, 4, 5, or 6 times higher or lower, respectively, than the control or standard. The difference in levels of markers or polynucleotides may be a "statistically significant difference"

Nucleic Acid Methods/Assays

- 15 As noted herein an ovarian cancer may be detected based on the levels of polynucleotides encoding kallikrein polypeptides, and optionally CA125, in a sample. Techniques for detecting polynucleotides such as polymerase chain reaction (PCR) and hybridization assays are well known in the art.

- Nucleotide probes for use in the detection of nucleic acid sequences in samples may be constructed using conventional methods known in the art. Suitable probes may be based on nucleic acid sequences
- 20 encoding at least 5 sequential amino acids from regions of nucleic acids encoding kallikrein polypeptides, and optionally CA125, preferably they comprise 15 to 40 nucleotides. A nucleotide probe may be labeled with a detectable substance such as a radioactive label that provides for an adequate signal and has sufficient half-life such as ^{32}P , ^3H , ^{14}C or the like. Other detectable substances that may be used include antigens that are recognized by a specific labeled antibody, fluorescent compounds, enzymes, antibodies specific for a
- 25 labeled antigen, and luminescent compounds. An appropriate label may be selected having regard to the rate of hybridization and binding of the probe to the nucleotide to be detected and the amount of nucleotide available for hybridization. Labeled probes may be hybridized to nucleic acids on solid supports such as nitrocellulose filters or nylon membranes as generally described in Sambrook et al, 1989, Molecular Cloning, A Laboratory Manual (2nd ed.). The nucleic acid probes may be used to detect polynucleotides
- 30 encoding kallikrein polypeptides, and optionally CA125, preferably in human cells. The nucleotide probes may also be useful in the diagnosis of ovarian cancer involving polynucleotides encoding kallikrein polypeptides, and optionally CA125, in monitoring the progression of such disorder; or monitoring a therapeutic treatment.

- Probes may be used in hybridization techniques to detect nucleic acids encoding a plurality of
- 35 kallikrein polypeptides, and optionally CA125. The technique generally involves contacting and incubating nucleic acids (e.g. recombinant DNA molecules, cloned genes) obtained from a sample from a patient or other cellular source with probes under conditions favorable for the specific annealing of the probes to complementary sequences in the nucleic acids. After incubation, the non-annealed nucleic acids are removed, and the presence of nucleic acids that have hybridized to the probe if any are detected.

- 14 -

The detection of polynucleotides encoding kallikrein polypeptides and optionally CA125, may involve the amplification of specific gene sequences using an amplification method such as polymerase chain reaction (PCR), followed by the analysis of the amplified molecules using techniques known to those skilled in the art. Suitable primers can be routinely designed by one of skill in the art.

5 By way of example, oligonucleotide primers may be employed in a PCR based assay to amplify a portion of nucleic acids encoding each of a plurality of kallikrein polypeptides, and optionally CA125, derived from a sample, wherein the oligonucleotide primers are specific for (i.e. hybridize to) polynucleotides encoding each of the plurality of kallikrein polypeptides, and optionally CA125. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel
10 electrophoresis.

In order to maximize hybridization under assay conditions, primers and probes employed in the methods of the invention generally have at least about 60%, preferably at least about 75% and more preferably at least about 90% identity to a portion of polynucleotides encoding a plurality of kallikrein polypeptides, and CA125. The primers and probes may be at least 10 nucleotides, and preferably at least 20
15 nucleotides in length. In an embodiment the primers and probes are at least about 10-40 nucleotides in length.

Hybridization and amplification techniques described herein may be used to assay qualitative and quantitative aspects of expression of polynucleotides encoding kallikrein polypeptides, and optionally CA125. For example, RNA may be isolated from a cell type or tissue known to express these
20 polynucleotides and tested utilizing the hybridization (e.g. standard Northern analyses) or PCR techniques referred to herein.

The primers and probes may be used in the above-described methods *in situ* i.e directly on tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections.

In an aspect of the invention, a method is provided employing reverse transcriptase-polymerase
25 chain reaction (RT-PCR), in which PCR is applied in combination with reverse transcription. Generally, RNA is extracted from a sample tissue using standard techniques (for example, guanidine isothiocyanate extraction as described by Chomcynski and Sacchi, Anal. Biochem. 162:156-159, 1987) and is reverse transcribed to produce cDNA. The cDNA is used as a template for a polymerase chain reaction. The cDNA is hybridized to sets of primers specifically designed against each of a plurality of kallikrein polynucleotide
30 sequences, and optionally CA125. Once the primer and template have annealed a DNA polymerase is employed to extend from the primer, to synthesize a copy of the template. The DNA strands are denatured, and the procedure is repeated many times until sufficient DNA is generated to allow visualization by ethidium bromide staining and agarose gel electrophoresis.

Amplification may be performed on samples obtained from a subject with suspected ovarian cancer
35 and an individual who is not afflicted with ovarian cancer. The reaction may be performed on several dilutions of cDNA spanning at least two orders of magnitude. A statistically significant difference in expression in several dilutions of the subject sample as compared to the same dilutions of the non-cancerous sample may be considered positive for the presence of ovarian cancer.

Oligonucleotides or longer fragments derived from polynucleotides encoding each of a plurality of

kallikrein polypeptides and optionally CA125, may be used as targets in a microarray. The microarray can be used to simultaneously monitor the expression levels of large numbers of genes. The information from the microarray may be used to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

5 The preparation, use, and analysis of microarrays are well known to a person skilled in the art. (See, for example, Brennan, T. M. et al. (1995) U.S. Pat. No. 5,474,796; Schena, et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995), PCT Application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R. A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M. J. et al. (1997) U.S. Pat. No. 5,605,662.)

10 Thus, the invention also includes an array comprising a plurality of polynucleotides encoding kallikrein marker(s), and optionally CA125 polynucleotides. The array can be used to assay expression of kallikrein polynucleotides, and optionally CA125 polynucleotides in the array. The invention allows the quantitation of expression of a plurality of kallikrein polynucleotides, and optionally CA125 polynucleotides.

15 In an embodiment, the array can be used to monitor the time course of expression of a plurality of kallikrein polynucleotides, and optionally CA125 polynucleotides, in the array. This can occur in various biological contexts such as tumor progression.

 The array is also useful for ascertaining differential expression patterns of a plurality of kallikrein polynucleotides and optionally CA125 polynucleotides, in normal and abnormal cells. This provides a battery of polynucleotides that could serve as molecular targets for diagnosis or therapeutic intervention.

20 **Protein Methods**

 Binding agents specific for a plurality of kallikrein markers and CA125 may be used for a variety of diagnostic and assay applications. There are a variety of assay formats known to the skilled artisan for using a binding agent to detect a target molecule in a sample. (For example, see Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988). In general, the presence or absence of an ovarian cancer in a subject may be determined by (a) contacting a sample from the subject with binding agents for a plurality of kallikrein polypeptides, and optionally CA125; (b) detecting in the sample levels of polypeptides that bind to the binding agents; and (c) comparing the levels of polypeptides with a predetermined standard or cut-off value.

30 “Binding agent” refers to a substance such as a polypeptide or antibody that specifically binds to a kallikrein or CA125 polypeptide. A substance “specifically binds” to a polypeptide if it reacts at a detectable level with the kallikrein or CA125 polypeptide, and does not react detectably with peptides containing unrelated sequences or sequences of different polypeptides. Binding properties may be assessed using an ELISA, which may be readily performed by those skilled in the art (see for example, Newton et al., Develop. Dynamics 197: 1-13, 1993).

35 A binding agent may be a ribosome, with or without a peptide component, an aptamer, an RNA molecule, or a polypeptide. A binding agent may be a polypeptide that comprises a kallikrein polypeptide or CA125 polypeptide sequence, a peptide variant thereof, or a non-peptide mimetic of such a sequence. By way of example a kallikrein polypeptide sequence may be a peptide portion of a kallikrein polypeptide that is capable of modulating a function mediated by the kallikrein polypeptide.

An aptamer includes a DNA or RNA molecule that binds to polynucleotides and polypeptides. An aptamer that binds to a polypeptide (or binding domain) of a kallikrein polypeptide or a polynucleotide encoding a kallikrein polypeptide can be produced using conventional techniques, without undue experimentation. [For example, see the following publications describing *in vitro* selection of aptamers: Klug et al., Mol. Biol. Reports 20:97-107 (1994); Wallis et al., Chem. Biol. 2:543-552 (1995); Ellington, Curr. Biol. 4:427-429 (1994); Lato et al., Chem. Biol. 2:291-303 (1995); Conrad et al., Mol. Div. 1:69-78 (1995); and Uphoff et al., Curr. Opin. Struct. Biol. 6:281-287 (1996)].

In certain other preferred embodiments, the binding agent is an antibody.

In an aspect the present invention provides a diagnostic method for monitoring or diagnosing ovarian cancer in a subject by quantitating a plurality of kallikrein polypeptides, and optionally CA125, in a biological sample from the subject comprising reacting the sample with antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, which are directly or indirectly labelled with detectable substances, and detecting the detectable substances.

In an aspect of the invention, a method for detecting ovarian cancer is provided comprising:

- (a) obtaining a sample suspected of containing a plurality of kallikrein polypeptides, and optionally CA125, wherein the kallikrein polypeptides comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11;
- (b) contacting the sample with antibodies that specifically bind to the plurality of kallikrein polypeptides, and optionally CA125, under conditions effective to bind the antibodies and form complexes;
- (c) measuring the amount of kallikrein polypeptides, and optionally CA125, present in the sample by quantitating the amount of the complexes; and
- (d) comparing the amount of kallikrein polypeptides, and optionally CA125, present in the samples with the amount of polypeptides in a control, wherein a change or significant difference in the amount of polypeptides in the sample compared with the amount in the control is indicative of ovarian cancer.

In an embodiment, the invention contemplates a method for monitoring the progression of ovarian cancer in an individual, comprising:

- (a) contacting antibodies which bind to each of a plurality of kallikrein polypeptides, and optionally CA125, with a sample from the individual so as to form binary complexes comprising each of the antibodies and polypeptides in the sample;
- (b) determining or detecting the presence or amount of complex formation in the sample;
- (c) repeating steps (a) and (b) at a point later in time; and
- (d) comparing the result of step (b) with the result of step (c), wherein a difference in the amount of complex formation is indicative of the stage and/or progression of the ovarian cancer in said individual.

The amount of complexes may also be compared to a value representative of the amount of the complexes from an individual not at risk of, or afflicted with, ovarian cancer at different stages.

Thus, antibodies specifically reactive with each of a plurality of kallikrein polypeptides, and CA125, or derivatives, such as enzyme conjugates or labeled derivatives, may be used to detect a plurality of kallikrein polypeptides, and optionally CA125, in various samples (e.g. biological materials). They may be used as diagnostic or prognostic reagents and they may be used to detect abnormalities in the levels of expression of a plurality of kallikrein polypeptides, and optionally CA125, or abnormalities in the structure, and/or temporal, tissue, cellular, or subcellular location of a plurality of kallikrein polypeptides, and optionally CA125. Antibodies may also be used to screen potentially therapeutic compounds *in vitro* to determine their effects on ovarian cancer involving a plurality of kallikrein polypeptides, and optionally CA125, and other conditions. *In vitro* immunoassays may also be used to assess or monitor the efficacy of particular therapies.

Antibodies may be used in any known immunoassays that rely on the binding interaction between antigenic determinants of a plurality of kallikrein polypeptides, and optionally CA125, and the antibodies. Examples of such assays are radioimmunoassays, enzyme immunoassays (e.g. ELISA), immunofluorescence, immunoprecipitation, latex agglutination, hemagglutination, and histochemical tests. These terms are well understood by those skilled in the art. A person skilled in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

In particular, the antibodies may be used in immunohistochemical analyses, for example, at the cellular and sub-subcellular level, to detect a plurality of kallikrein polypeptides, and optionally CA125, to localize them to particular ovarian tumor cells and tissues, and to specific subcellular locations, and to quantitate the level of expression.

Antibodies for use in the present invention include monoclonal or polyclonal antibodies, immunologically active fragments (e.g. a Fab or (Fab)₂ fragments), antibody heavy chains, humanized antibodies, antibody light chains, genetically engineered single chain F_v molecules (Ladner et al, U.S. Pat. No. 4,946,778), chimeric antibodies, for example, antibodies which contain the binding specificity of murine antibodies, but in which the remaining portions are of human origin, or derivatives, such as enzyme conjugates or labeled derivatives.

Antibodies including monoclonal and polyclonal antibodies, fragments and chimeras, may be prepared using methods known to those skilled in the art. Isolated native or recombinant kallikrein polypeptides or CA125 may be utilized to prepare antibodies. See, for example, Kohler et al. (1975) Nature 256:495-497; Kozbor et al. (1985) J. Immunol Methods 81:31-42; Cote et al. (1983) Proc Natl Acad Sci 80:2026-2030; and Cole et al. (1984) Mol Cell Biol 62:109-120 for the preparation of monoclonal antibodies; Huse et al. (1989) Science 246:1275-1281 for the preparation of monoclonal Fab fragments; and, Pound (1998) Immunochemical Protocols, Humana Press, Totowa, N.J for the preparation of phagemid or B-lymphocyte immunoglobulin libraries to identify antibodies. The antibodies specific for kallikrein polypeptides or CA125 used in the methods of the invention may also be obtained from scientific or commercial sources.

In an embodiment of the invention, antibodies are reactive against kallikrein polypeptides or CA125 if they bind with a K_a of greater than or equal to 10⁻⁷ M.

Antibodies that bind to kallikrein polypeptides or CA125 may be labelled with a detectable

substance and localised in biological samples based upon the presence of the detectable substance. Examples of detectable substances include, but are not limited to, the following: radioisotopes (e.g., ^3H , ^{14}C , ^{35}S , ^{125}I , ^{131}I), fluorescent labels (e.g., FITC, rhodamine, lanthanide phosphors), luminescent labels such as luminol, enzymatic labels (e.g., horseradish peroxidase, beta-galactosidase, luciferase, alkaline phosphatase, acetylcholinesterase), biotinyl groups (which can be detected by marked avidin e.g., streptavidin containing a fluorescent marker or enzymatic activity that can be detected by optical or colorimetric methods), and predetermined polypeptide epitopes recognized by a secondary reporter (e.g., leucine zipper pair sequences, binding sites for secondary antibodies, metal binding domains, epitope tags). In some embodiments, labels are attached via spacer arms of various lengths to reduce potential steric hindrance. Antibodies may also be coupled to electron dense substances, such as ferritin or colloidal gold, which are readily visualised by electron microscopy.

Indirect methods may also be employed in which the primary antigen-antibody reaction is amplified by the introduction of a second antibody, having specificity for the antibody reactive against a kallikrein polypeptide or CA125. The second antibody may be labeled with a detectable substance to detect the primary antigen-antibody reaction. By way of example, if the antibody having specificity against a kallikrein polypeptide is a rabbit IgG antibody, the second antibody may be goat anti-rabbit gamma-globulin labelled with a detectable substance as described herein.

Methods for conjugating or labelling the antibodies discussed above may be readily accomplished by one of ordinary skill in the art. (See for example Inman, *Methods In Enzymology*, Vol. 34, Affinity Techniques, Enzyme Purification: Part B, Jakoby and Wichek (eds.), Academic Press, New York, p. 30, 1974; and Wilchek and Bayer, "The Avidin-Biotin Complex in Bioanalytical Applications," *Anal. Biochem.* 171:1-32, 1988 re methods for conjugating or labelling the antibodies with enzyme or ligand binding partner).

Cytochemical techniques known in the art for localizing antigens using light and electron microscopy may be used to detect a plurality of kallikrein polypeptides, and optionally CA125. Generally, antibodies may be labeled with detectable substances and kallikrein polypeptides, and optionally CA125, may be localised in tissues and cells based upon the presence of the detectable substance.

In the context of the methods of the invention, the sample, binding agents (e.g. antibodies) for a plurality of kallikrein polypeptides, and CA125 may be immobilized on a carrier or support. Examples of suitable carriers or supports are agarose, cellulose, nitrocellulose, dextran, Sephadex, Sepharose, liposomes, carboxymethyl cellulose, polyacrylamides, polystyrene, gabbros, filter paper, magnetite, ion-exchange resin, plastic film, plastic tube, glass, polyamine-methyl vinyl-ether-maleic acid copolymer, amino acid copolymer, ethylene-maleic acid copolymer, nylon, silk, etc. The support material may have any possible configuration including spherical (e.g. bead), cylindrical (e.g. inside surface of a test tube or well, or the external surface of a rod), or flat (e.g. sheet, test strip). Thus, the carrier may be in the shape of, for example, a tube, test plate, well, beads, disc, sphere, etc. The immobilized material may be prepared by reacting the material with a suitable insoluble carrier using known chemical or physical methods, for example, cyanogen bromide coupling. Binding agents (e.g. antibodies) may be indirectly immobilized using second binding agents specific for the first binding agent. For example, mouse antibodies specific for a kallikrein polypeptide may

be immobilized using sheep anti-mouse IgG Fc fragment specific antibody coated on the carrier or support.

Where radioactive labels are used as a detectable substance, a plurality of kallikrein polypeptides, and optionally CA125, may be localized by radioautography. The results of radioautography may be quantitated by determining the density of particles in the radioautographs by various optical methods, or by counting the grains.

Time-resolved fluorometry may be used to detect a signal. For example, the method described in Christopoulos TK and Diamandis EP Anal Chem 1992:64:342-346 may be used with a conventional time-resolved fluorometer.

Therefore, in accordance with an embodiment of the invention, a method is provided wherein antibodies specific for each of a plurality of kallikrein polypeptides, and optionally CA125, are labelled with enzymes, substrates for the enzymes are added wherein the substrates are selected so that the substrates, or a reaction product of the enzymes and substrates, form fluorescent complexes with lanthanide metals. Lanthanide metals are added and the plurality of kallikrein polypeptides, and optionally CA125, are quantitated in the sample by measuring fluorescence of the fluorescent complexes. Antibodies specific for CA125 and each of a plurality of kallikrein polypeptides may be directly or indirectly labelled with enzymes. Enzymes are selected based on the ability of a substrate of the enzyme, or a reaction product of the enzyme and substrate, to complex with lanthanide metals such as europium and terbium. Examples of suitable enzymes include alkaline phosphatase and β -galactosidase.

Examples of enzymes and substrates for enzymes that provide such fluorescent complexes are described in U.S. Patent No. 5,312,922 to Diamandis. By way of example, when the antibody is directly or indirectly labelled with alkaline phosphatase the substrate employed in the method may be 4-methylumbelliferyl phosphate, 5-fluorosalicyl phosphate, or diflunisal phosphate. The fluorescence intensity of the complexes is typically measured using a time-resolved fluorometer e.g. a CyberFluor 615 Imunoanalyzer (Nordion International, Kanata, Ontario).

Antibodies specific for a plurality of kallikrein polypeptides and CA125 may also be indirectly labelled with enzymes. For example, an antibody may be conjugated to one partner of a ligand binding pair, and the enzyme may be coupled to the other partner of the ligand binding pair. Representative examples include avidin-biotin, and riboflavin-riboflavin binding protein. In another embodiment, antibodies specific for the anti-kallikrein antibodies or anti-CA125 antibodies are labeled with an enzyme.

In accordance with an embodiment, the present invention provides means for determining a plurality of kallikrein polypeptides, and optionally CA125, in a sample, in particular a serum sample, by measuring a plurality of kallikrein polypeptides, and optionally CA125, by immunoassay. It will be evident to a skilled artisan that a variety of immunoassay methods can be used to measure a plurality of kallikrein polypeptides and CA125 in serum. In general, an immunoassay method may be competitive or noncompetitive. Competitive methods typically employ immobilized or immobilizable antibodies to each of a plurality of kallikrein polypeptides, and optionally CA125, and a labeled form of each of a plurality of kallikrein polypeptides, and optionally CA125. Kallikrein polypeptides and CA125 and labeled kallikrein polypeptides and CA125 compete for binding to anti-kallikrein antibodies and anti-CA125 antibodies. After separation of the resulting labeled kallikrein polypeptides and CA125 that have become bound to anti-

kallikrein polypeptides and anti- CA125 (bound fraction) from that which has remained unbound (unbound fraction), the amount of the label in either bound or unbound fraction is measured and may be correlated with the amount of kallikrein polypeptides, and optionally CA125, in the test sample in any conventional manner, e.g., by comparison to a standard curve.

5 In an aspect, a non-competitive method is used for the determination of a plurality of kallikrein polypeptides, and optionally CA125, with the most common method being the "sandwich" method. In this assay, two types of antibodies specific for each of a plurality of kallikrein polypeptides, and optionally CA125 are employed. One type of antibody is directly or indirectly labeled (sometimes referred to as the "detection antibody") and the other is immobilized or immobilizable (sometimes referred to as the "capture antibody"). The capture and detection antibodies can be contacted simultaneously or sequentially with a test sample. Sequential methods can be accomplished by incubating capture antibodies with the sample, and adding the detection antibodies at a predetermined time thereafter (sometimes referred to as the "forward" method); or the detection antibodies can be incubated with the sample first and then the capture antibodies added (sometimes referred to as the "reverse" method). After the necessary incubation(s) have occurred, to 10 complete the assay, the capture antibodies are separated from the liquid test mixture, and labels are measured in at least a portion of the separated capture antibody phase or the remainder of the liquid test mixture. Generally the labels are measured in the capture antibody phase since it comprises kallikrein polypeptides, and optionally CA125, bound by ("sandwiched" between) the capture and detection antibodies. In an embodiment, the label may be measured without separating the capture antibodies and liquid test mixture.

20 In a typical two-site immunometric assay for a plurality of kallikrein polypeptides, and optionally CA125, one or both of the capture and detection antibodies are polyclonal antibodies or one or both of the capture and detection antibodies are monoclonal antibodies (i.e. polyclonal/polyclonal, monoclonal/monoclonal, or monoclonal/polyclonal). The labels used with the detection antibodies can be selected from any of those known conventionally in the art. The labels may be an enzyme or a chemiluminescent moiety, but it can also be a radioactive isotope, a fluorophor, a detectable ligand (e.g., 25 detectable by a secondary binding by a labeled binding partner for the ligand), and the like. Preferably antibodies are labelled with enzymes which are detected by adding substrates that are selected so that a reaction product of the enzymes and substrates forms fluorescent complexes. Capture antibodies may be selected so that they provide a means for being separated from the remainder of the test mixture. Accordingly, the capture antibodies can be introduced to the assay in an already immobilized or insoluble form, or can be in an immobilizable form, that is, a form which enables immobilization to be accomplished subsequent to introduction of the capture antibodies to the assay. An immobilized capture antibody may comprise an antibody covalently or noncovalently attached to a solid phase such as a magnetic particle, a latex particle, a microtiter plate well, a bead, a cuvette, or other reaction vessel. An example of an 30 immobilizable capture antibody is antibody which has been chemically modified with a ligand moiety, e.g., a hapten, biotin, or the like, and which can be subsequently immobilized by contact with an immobilized form of a binding partner for the ligand, e.g., an antibody, avidin, or the like. In an embodiment, a capture antibody may be immobilized using a species specific antibody for the capture antibody that is bound to the solid phase.

A particular sandwich immunoassay method of the invention employs two types of antibodies, first antibodies are reactive against each of a plurality of kallikrein polypeptides, and optionally CA125, and second antibodies having specificity against antibodies reactive against each of a plurality of kallikrein polypeptides, and optionally CA125, labelled with enzymatic labels, and fluorogenic substrates for the enzymes. An enzyme may be alkaline phosphatase (ALP) and the substrate is 5-fluorosalicyl phosphate. ALP cleaves phosphate out of the fluorogenic substrate, 5-fluorosalicyl phosphate, to produce 5-fluorosalicylic acid (FSA). 5-Fluorosalicylic acid can then form a highly fluorescent ternary complex of the form FSA-Tb(3+)-EDTA, which can be quantified by measuring the Tb3+ fluorescence in a time-resolved mode. Fluorescence intensity is measured using a time-resolved fluorometer as described herein.

The above-described immunoassay methods and formats are intended to be exemplary and are not limiting.

Computer Systems

Computer readable media comprising a plurality of kallikrein markers, and optionally CA125, is also provided. "Computer readable media" refers to any medium that can be read and accessed directly by a computer, including but not limited to magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. Thus, the invention contemplates computer readable medium having recorded thereon markers identified for patients and controls.

"Recorded" refers to a process for storing information on computer readable medium. The skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising information on a plurality of kallikrein markers, and optionally CA125.

A variety of data processor programs and formats can be used to store information on a plurality of kallikrein markers, and optionally CA125, on computer readable medium. For example, the information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. Any number of dataprocessor structuring formats (e.g., text file or database) may be adapted in order to obtain computer readable medium having recorded thereon the marker information.

By providing the marker information in computer readable form, one can routinely access the information for a variety of purposes. For example, one skilled in the art can use the information in computer readable form to compare marker information obtained during or following therapy with the information stored within the data storage means.

The invention provides a medium for holding instructions for performing a method for determining whether a patient has ovarian cancer or a pre-disposition to ovarian cancer, comprising determining the presence or absence of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, and based on the presence or absence of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, determining whether the patient has ovarian cancer or a pre-disposition to

ovarian cancer, and optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention also provides in an electronic system and/or in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a plurality of kallikrein markers, and optionally CA125, and/or polynucleotides encoding same, comprising determining the presence or absence of a plurality of kallikrein markers, and optionally CA125, and/or polynucleotides encoding same, and based on the presence or absence of the plurality of kallikrein markers, and optionally CA125, and/or polynucleotides encoding same, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer, and optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention further provides in a network, a method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a plurality of kallikrein markers, optionally CA125 and/or polynucleotides encoding same, comprising: (a) receiving phenotypic information on the subject and information on a plurality of kallikrein markers, optionally CA125 and/or polynucleotides encoding same associated with samples from the subject; (b) acquiring information from the network corresponding to the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same; and (c) based on the phenotypic information and information on the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer; and (d) optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

The invention still further provides a system for identifying selected records that identify an ovarian cancer cell. A system of the invention generally comprises a digital computer; a database server coupled to the computer; a database coupled to the database server having data stored therein, the data comprising records of data comprising a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, and a code mechanism for applying queries based upon a desired selection criteria to the data file in the database to produce reports of records which match the desired selection criteria.

In an aspect of the invention a method is provided for detecting an ovarian cancer cell using a computer having a processor, memory, display, and input/output devices, the method comprising the steps of:

- (a) creating records of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, isolated from a sample suspected of containing an ovarian cancer cell;
- (b) providing a database comprising records of data comprising a plurality of kallikrein markers, optionally CA125, wherein the markers are kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and/or comprising polynucleotides encoding same; and
- (c) using a code mechanism for applying queries based upon a desired selection criteria to the data file in the database to produce reports of records of step (a) which provide a match of the desired selection criteria of the database of step (b) the presence of a match being a

- 23 -

positive indication that the markers of step (a) have been isolated from a cell that is an ovarian cancer cell.

The invention contemplates a business method for determining whether a subject has ovarian cancer or a pre-disposition to ovarian cancer associated with a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, comprising: (a) receiving phenotypic information on the subject and information on a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, associated with samples from the subject; (b) acquiring information from a network corresponding to the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same; and (c) based on the phenotypic information, information on a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, and acquired information, determining whether the subject has ovarian cancer or a pre-disposition to ovarian cancer; and (d) optionally recommending treatment for the ovarian cancer or pre-ovarian cancer condition.

Imaging Methods

Antibodies specific for each of a plurality of kallikrein polypeptides, and optionally CA125, may also be used in imaging methodologies in the management of ovarian cancer. The invention provides a method for imaging tumors associated with a plurality of kallikrein polypeptides, and optionally CA125.

In an embodiment the method is an *in vivo* method and a subject or patient is administered imaging agents that carry imaging labels and are capable of targeting or binding to each of a plurality of kallikrein polypeptides, and optionally CA125. In the method each imaging agent is labeled so that it can be distinguished during the imaging. The imaging agents are allowed to incubate *in vivo* and bind to the plurality of kallikrein polypeptides, and optionally CA125, associated with ovarian tumors. The presence of label is localized to the ovarian cancer, and the localized label is detected using imaging devices known to those skilled in the art.

The imaging agents may be antibodies or chemical entities that recognize the plurality of kallikrein polypeptides, and optionally CA125. In an aspect of the invention an imaging agent is a polyclonal antibody or monoclonal antibody, or fragments thereof, or constructs thereof including but not limited to, single chain antibodies, bifunctional antibodies, molecular recognition units, and peptides or entities that mimic peptides. The antibodies specific for kallikrein polypeptides and CA125 used in the methods of the invention may be obtained from scientific or commercial sources, or isolated native or recombinant kallikrein and CA125 polypeptides may be utilized to prepare antibodies etc as described herein.

An imaging agent may be a peptide that mimics the epitope for an antibody specific for kallikrein polypeptide or CA125 and binds to kallikrein polypeptide or CA125. The peptide may be produced on a commercial synthesizer using conventional solid phase chemistry. By way of example, a peptide may be prepared that includes either tyrosine, lysine, or phenylalanine to which N_2S_2 chelate is complexed (See U.S. Patent No. 4,897,255). The anti-kallikrein peptide conjugate is then combined with a radiolabel (e.g. sodium ^{99m}Tc pertechnetate or sodium ^{188}Re perrhenate) and it may be used to locate a tumor producing a plurality of kallikrein polypeptides, and optionally CA125.

Imaging agents carry labels to image the plurality of kallikrein polypeptides and CA125. Agents may be labelled for use in radionuclide imaging. In particular, agents may be directly or indirectly labelled

with a radioisotope. Examples of radioisotopes that may be used in the present invention are the following:
²⁷⁷Ac, ²¹¹At, ¹²⁸Ba, ¹³¹Ba, ⁷Be, ²⁰⁴Bi, ²⁰⁵Bi, ²⁰⁶Bi, ⁷⁶Br, ⁷⁷Br, ⁸²Br, ¹⁰⁹Cd, ⁴⁷Ca, ¹¹C, ¹⁴C, ³⁶Cl, ⁴⁸Cr, ⁵¹Cr, ⁶²Cu,
⁶⁴Cu, ⁶⁷Cu, ¹⁶⁵Dy, ¹⁵⁵Eu, ¹⁸F, ¹⁵³Gd, ⁶⁶Ga, ⁶⁷Ga, ⁶⁸Ga, ⁷²Ga, ¹⁹⁸Au, ³H, ¹⁶⁶Ho, ¹¹¹In, ^{113m}In, ^{115m}In, ¹²³I, ¹²⁵I, ¹³¹I,
¹⁸⁹Ir, ^{191m}Ir, ¹⁹²Ir, ¹⁹⁴Ir, ⁵²Fe, ⁵⁵Fe, ⁵⁹Fe, ¹⁷⁷Lu, ¹⁵O, ^{191m-191}Os, ¹⁰⁹Pd, ³²P, ³³P, ⁴²K, ²²⁶Ra, ¹⁸⁶Re, ¹⁸⁸Re, ^{82m}Rb,
5 ¹⁵³Sm, ⁴⁶Sc, ⁴⁷Sc, ⁷²Se, ⁷⁵Se, ¹⁰⁵Ag, ²²Na, ²⁴Na, ⁸⁹Sr, ³⁵S, ³⁸S, ¹⁷⁷Ta, ⁹⁶Tc, ^{99m}Tc, ²⁰¹Tl, ²⁰²Tl, ¹¹³Sn, ^{117m}Sn,
¹²¹Sn, ¹⁶⁶Yb, ¹⁶⁹Yb, ¹⁷⁵Yb, ⁸⁸Y, ⁹⁰Y, ⁶²Zn and ⁶⁵Zn. Preferably the radioisotope is ¹³¹I, ¹²⁵I, ¹²³I, ¹¹¹I, ^{99m}Tc,
⁹⁰Y, ¹⁸⁶Re, ¹⁸⁸Re, ³²P, ¹⁵³Sm, ⁶⁷Ga, ²⁰¹Tl ⁷⁷Br, or ¹⁸F, and it is imaged with a photoscanning device.

Procedures for labeling biological agents with the radioactive isotopes are generally known in the
art. U.S. Pat. No. 4,302,438 describes tritium labeling procedures. Procedures for iodinating, tritium labeling,
10 and ³⁵S labeling especially adapted for murine monoclonal antibodies are described by Goding, J. W. (supra,
pp 124-126) and the references cited therein. Other procedures for iodinating biological agents, such as
antibodies, binding portions thereof, probes, or ligands, are described in the scientific literature (see Hunter
and Greenwood, Nature 144:945 (1962), David et al., Biochemistry 13:1014-1021 (1974), and U.S. Pat. Nos.
3,867,517 and 4,376,110). Iodinating procedures for agents are described by Greenwood, F. et al., Biochem.
15 J. 89:114-123 (1963); Marchalonis, J., Biochem. J. 113:299-305 (1969); and Morrison, M. et al.,
Immunochemistry, 289-297 (1971). ^{99m}Tc-labeling procedures are described by Rhodes, B. et al. in
Burchiel, S. et al. (eds.), Tumor Imaging: The Radioimmunochemical Detection of Cancer, New York:
Masson 111-123 (1982) and the references cited therein. Labelling of antibodies or fragments with
technetium-99m are also described for example in U.S. Pat. No. 5,317,091, U.S. Pat. No. 4,478,815, U.S.
20 Pat. No. 4,478,818, U.S. Pat. No. 4,472,371, U.S. Pat. No. Re 32,417, and U.S. Pat. No. 4,311,688.
Procedures suitable for ¹¹¹In-labeling biological agents are described by Hnatowich, D. J. et al., J. Immul.
Methods, 65:147-157 (1983), Hnatowich, D. et al., J. Applied Radiation, 35:554-557 (1984), and Buckley,
R. G. et al., F.E.B.S. 166:202-204 (1984).

An imaging agent may also be labeled with a paramagnetic isotope for purposes of an *in vivo*
25 method of the invention. Examples of elements that are useful in magnetic resonance imaging include
gadolinium, terbium, tin, iron, or isotopes thereof. (See, for example, Schaefer et al., (1989) JACC 14, 472-
480; Shreve et al., (1986) Magn. Reson. Med. 3, 336-340; Wolf, G L., (1984) Physiol. Chem. Phys. Med.
NMR 16, 93-95; Wesbey et al., (1984) Physiol. Chem. Phys. Med. NMR 16, 145-155; Runge et al., (1984)
Invest. Radiol. 19, 408-415 for discussions on *in vivo* nuclear magnetic resonance imaging.)

In the case of radiolabeled agents, the agents may be administered to the patient, localized to the
tumor having a plurality of kallikrein polypeptides, and optionally CA125, with which the agents bind, and
detected or "imaged" *in vivo* using known techniques such as radionuclear scanning using, for example, a
gamma camera or emission tomography. [See for example, A. R. Bradwell et al., "Developments in
Antibody Imaging", Monoclonal Antibodies for Cancer Detection and Therapy, R. W. Baldwin et al., (eds.),
35 pp. 65-85 (Academic Press 1985)]. A positron emission transaxial tomography scanner, such as designated
Pet VI located at Brookhaven National Laboratory, can also be used where the radiolabel emits positrons
(e.g., ¹¹C, ¹⁸F, ¹⁵O, and ¹³N).

Whole body imaging techniques using radioisotope labeled agents can be used for locating both
primary tumors and tumors which have metastasized. Antibodies specific for a plurality of kallikrein

polypeptides, and optionally CA125, or fragments thereof having the same epitope specificity, are bound to a suitable radioisotope, or a combination thereof, and administered parenterally. For ovarian cancer, administration preferably is intravenous. The bio-distribution of the labels can be monitored by scintigraphy, and accumulations of the labels can be related to the presence of ovarian cancer cells. Whole body imaging techniques are described in U.S. Pat. Nos. 4,036,945 and 4,311,688. Other examples of agents useful for diagnosis and therapeutic use that can be coupled to antibodies and antibody fragments include metallothionein and fragments (see, U.S. Pat. No. 4,732,864). These agents are useful in diagnosis, staging and visualization of cancer, in particular ovarian cancer, so that surgical and/or radiation treatment protocols can be used more efficiently.

10 Screening Methods

The invention also contemplates methods for evaluating test agents or compounds for their ability to inhibit ovarian cancer or potentially contribute to ovarian cancer. Test agents and compounds include but are not limited to peptides such as soluble peptides including Ig-tailed fusion peptides, members of random peptide libraries and combinatorial chemistry-derived molecular libraries made of D- and/or L-configuration amino acids, phosphopeptides (including members of random or partially degenerate, directed phosphopeptide libraries), antibodies [e.g. polyclonal, monoclonal, humanized, anti-idiotypic, chimeric, single chain antibodies, fragments, (e.g. Fab, F(ab)₂, and Fab expression library fragments, and epitope-binding fragments thereof)], nucleic acids (e.g. antisense, interference RNA) and small organic or inorganic molecules. The agents or compounds may be endogenous physiological compounds or natural or synthetic compounds.

The invention also provides a method for assessing the potential efficacy of a test agent for inhibiting ovarian cancer in a patient, the method comprising comparing:

- (a) levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a first sample obtained from a patient and exposed to the test agent, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and
- (b) levels of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a second sample obtained from the patient, wherein the sample is not exposed to the test agent, wherein a significant difference in the levels of expression of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the first sample, relative to the second sample, is an indication that the test agent is potentially efficacious for inhibiting ovarian cancer in the patient.

The first and second samples may be portions of a single sample obtained from a patient or portions of pooled samples obtained from a patient.

In an aspect, the invention provides a method of selecting an agent for inhibiting ovarian cancer in a patient comprising:

- (a) obtaining a sample comprising cancer cells from the patient;
- (b) separately maintaining aliquots of the sample in the presence of a plurality of test agents;
- (c) comparing a plurality of kallikrein markers, optionally CA125, and/or polynucleotides

- 26 -

encoding same, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and

- 5 (d) selecting one of the test agents which alters the levels of the kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot containing that test agent, relative to other test agents.

Still another aspect of the present invention provides a method of conducting a drug discovery business comprising:

- 10 (a) providing one or more methods or assay systems for identifying agents that inhibit ovarian cancer in a patient;
- (b) conducting therapeutic profiling of agents identified in step (a), or further analogs thereof, for efficacy and toxicity in animals; and
- (c) formulating a pharmaceutical preparation including one or more agents identified in step (b) as having an acceptable therapeutic profile.

15 In certain embodiments, the subject method can also include a step of establishing a distribution system for distributing the pharmaceutical preparation for sale, and may optionally include establishing a sales group for marketing the pharmaceutical preparation.

The invention also contemplates a method of assessing the ovarian carcinogenic potential of a test compound comprising:

- 20 (a) maintaining separate aliquots of ovarian cells in the presence and absence of the test compound; and
- (b) comparing a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.

25 A significant difference between the levels of the markers in the aliquot maintained in the presence of (or exposed to) the test compound relative to the aliquot maintained in the absence of the test compound, indicates that the test compound possesses ovarian carcinogenic potential.

Kits

30 The methods described herein may be performed by utilizing pre-packaged diagnostic kits comprising at least a plurality of kallikrein nucleic acids or binding agents (e.g. antibodies) or CA125 nucleic acids or binding agents described herein, which may be conveniently used, e.g., in clinical settings, to screen and diagnose patients, and to screen and identify those individuals afflicted with or exhibiting a predisposition to ovarian cancer.

35 Thus, the invention also contemplates kits for carrying out the methods of the invention. Such kits typically comprise two or more components required for performing a diagnostic assay. Components include but are not limited to compounds, reagents, containers, and/or equipment.

In an embodiment, a container with a kit comprises binding agents as described herein. By way of example, the kit may contain antibodies specific for a plurality of kallikrein polypeptides, and optionally

CA125, antibodies against the antibodies labelled with enzymes; and substrates for the enzymes. The kit may also contain microtiter plate wells, standards, assay diluent, wash buffer, adhesive plate covers, and/or instructions for carrying out a method of the invention using the kit.

5 In an aspect of the invention, the kit includes antibodies or antibody fragments which bind specifically to epitopes of each of a plurality of kallikrein polypeptides, and optionally CA125, and means for detecting binding of the antibodies to epitopes associated with tumor cells, either as concentrates (including lyophilized compositions), which may be further diluted prior to use or at the concentration of use, where the vials may include one or more dosages. Where the kits are intended for *in vivo* use, single dosages may be provided in sterilized containers, having the desired amount and concentration of agents.

10 Containers that provide a formulation for direct use, usually do not require other reagents, as for example, where the kit contains radiolabelled antibody preparations for *in vivo* imaging.

A kit may be designed to detect the level of polynucleotides encoding kallikrein polypeptides, and optionally CA125 polynucleotides, in a sample. Such kits generally comprise oligonucleotide probes or primers, as described herein, that hybridize to a plurality of polynucleotides encoding kallikrein polypeptides

15 and optionally CA125. Such oligonucleotides may be used, for example, within a PCR or hybridization procedure. Additional components that may be present within the kits include second oligonucleotides and/or diagnostic reagents to facilitate detection of a plurality polynucleotides encoding kallikrein polypeptides, and optionally CA125 polynucleotides.

The reagents suitable for applying the screening methods of the invention to evaluate compounds

20 may be packaged into convenient kits described herein providing the necessary materials packaged into suitable containers.

Applications

Kallikrein polypeptides (in particular, kallikrein 5, 6, 10 and 11), optionally in combination with CA125, are targets for ovarian cancer immunotherapy. Such immunotherapeutic methods include the use of

25 antibody therapy, *in vivo* vaccines, and *ex vivo* immunotherapy approaches.

In one aspect, the invention provides antibodies specific for a plurality of kallikrein polypeptides (for example, kallikreins 5, 6, 10 and 11) and optionally CA125, that may be used systemically to treat ovarian cancer. Preferably antibodies are used that target the tumor cells but not the surrounding non-tumor cells and tissue. Thus, the invention provides a method of treating a patient susceptible to, or having a cancer

30 that expresses a plurality of kallikrein polypeptides, and optionally CA125, comprising administering to the patient an effective amount of antibodies that bind specifically to a plurality of kallikrein polypeptides, and optionally CA125. In another aspect, the invention provides a method of inhibiting the growth of tumor cells expressing a plurality of kallikrein polypeptides, and optionally CA125, comprising administering to a patient antibodies which bind specifically to the plurality of kallikrein polypeptides, and optionally CA125,

35 in amounts effective to inhibit growth of the tumor cells. Antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, may also be used in a method for selectively inhibiting the growth of, or killing a cell expressing a plurality of kallikrein polypeptides, and optionally CA125, comprising reacting antibody immunoconjugates or immunotoxins with the cell in an amount sufficient to inhibit the growth of, or kill the cell.

By way of example, unconjugated antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, may be introduced into a patient such that the antibodies bind to cancer cells expressing a plurality of kallikrein polypeptides, and optionally CA125, and mediate growth inhibition of such cells (including the destruction thereof), and the tumor, by mechanisms which may include complement-mediated
5 cytotoxicity, antibody-dependent cellular cytotoxicity, altering the physiologic function of a plurality of kallikrein polypeptides, and optionally CA125, and/or the inhibition of ligand binding or signal transduction pathways. In addition to unconjugated antibodies, antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, conjugated to therapeutic agents (e.g. immunoconjugates) may also be used therapeutically to deliver the agents directly to tumor cells expressing a plurality of kallikrein
10 polypeptides, and optionally CA125, and thereby destroy the tumor. Examples of such agents include abrin, ricin A, *Pseudomonas* exotoxin, or diphtheria toxin, proteins such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, and biological response modifiers such as lymphokines, interleukin-1, interleukin-2, interleukin-6, granulocyte macrophage colony stimulating factor, granulocyte colony stimulating factor, or other growth factors.

Cancer immunotherapy using antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, may utilize the various approaches that have been successfully employed for cancers, including but not limited to colon cancer (Arlen et al., 1998, Crit Rev Immunol 18: 133-138), multiple myeloma (Ozaki et al., 1997, Blood 90: 3179-3186; Tsunenati et al., 1997, Blood 90: 2437-2444), gastric cancer (Kasprzyk et al., 1992, Cancer Res 52: 2771-2776), B-cell lymphoma (Funakoshi et al., 1996, J
20 Immunther Emphasis Tumor Immunol 19: 93-101), leukemia (Zhong et al., 1996, Leuk Res 20: 581-589), colorectal cancer (Moun et al., 1994, Cancer Res 54: 6160-6166); Velders et al., 1995, Cancer Res 55: 4398-4403), and breast cancer (Shepard et al., 1991, J Clin Immunol 11: 117-127).

In the practice of a method of the invention, antibodies specific for a plurality of kallikrein polypeptides, optionally in combination with antibodies specific for CA125, capable of inhibiting the growth
25 of cancer cells expressing a plurality of kallikrein polypeptides, and optionally CA125, are administered in a therapeutically effective amount to cancer patients whose tumors express or overexpress a plurality of kallikrein polypeptides, and optionally CA125. The invention may provide a specific, effective and long-needed treatment for ovarian cancer. The antibody therapy methods of the invention may be combined with other therapies including chemotherapy and radiation.

Patients may be evaluated for the presence and levels of a plurality of kallikrein polypeptides, and optionally CA125, expression and overexpression in tumors, preferably using immunohistochemical assessments of tumor tissue, quantitative imaging as described herein, or other techniques capable of reliably indicating the presence and degree of expression of a plurality of kallikrein polypeptides, and optionally CA125. Immunohistochemical analysis of tumor biopsies or surgical specimens may be employed for this
35 purpose.

Antibodies specific for a plurality of kallikrein polypeptides and CA125 useful in treating cancer include those that are capable of initiating a potent immune response against the tumor and those that are capable of direct cytotoxicity. In this regard, the antibodies may elicit tumor cell lysis by either complement-mediated or antibody-dependent cell cytotoxicity (ADCC) mechanisms, both of which require an intact Fc

portion of the immunoglobulin molecule for interaction with effector cell Fc receptor sites or complement proteins. In addition, antibodies specific for a plurality of kallikrein polypeptides and CA125 that exert a direct biological effect on tumor growth are useful in the practice of the invention. Such antibodies may not require the complete immunoglobulin to exert the effect. Potential mechanisms by which such directly cytotoxic antibodies may act include inhibition of cell growth, modulation of cellular differentiation, modulation of tumor angiogenesis factor profiles, and the induction of apoptosis. The mechanism by which a particular antibody exerts an anti-tumor effect may be evaluated using any number of *in vitro* assays designed to determine ADCC, antibody-dependent macrophage-mediated cytotoxicity (ADMMC), complement-mediated cell lysis, and others known in the art.

The anti-tumor activity of a combination of antibodies specific for a plurality of kallikrein polypeptides and optionally CA125, may be evaluated *in vivo* using a suitable animal model. Xenogenic cancer models, wherein human cancer explants or passaged xenograft tissues are introduced into immune compromised animals, such as nude or SCID mice, may be employed.

The methods of the invention contemplate the administration of combinations, or "cocktails" of different individual antibodies recognizing epitopes of a plurality of kallikrein polypeptides, and optionally CA125. Such cocktails may have certain advantages inasmuch as they contain antibodies that bind to different epitopes and/or exploit different effector mechanisms or combine directly cytotoxic antibodies with antibodies that rely on immune effector functionality. Such antibodies in combination may exhibit synergistic therapeutic effects. In addition, the administration of the antibodies may be combined with other therapeutic agents, including but not limited to chemotherapeutic agents, androgen-blockers, and immune modulators (e.g., IL2, GM-CSF). The antibodies may be administered in their "naked" or unconjugated form, or may have therapeutic agents conjugated to them.

The antibodies specific for a plurality of kallikrein polypeptides and optionally CA125, used in the practice of the method of the invention may be formulated into pharmaceutical compositions comprising a carrier suitable for the desired delivery method. Suitable carriers include any material which when combined with the antibodies retains the anti-tumor function of the antibodies and is non-reactive with the subject's immune systems. Examples include any of a number of standard pharmaceutical carriers such as sterile phosphate buffered saline solutions, bacteriostatic water, and the like (see, generally, Remington's Pharmaceutical Sciences 16^{sup}.th Edition, A. Osal., Ed., 1980).

Antibody formulations may be administered via any route capable of delivering the antibodies to the tumor site. Routes of administration include, but are not limited to, intravenous, intraperitoneal, intramuscular, intratumor, intradermal, and the like. Preferably, the route of administration is by intravenous injection. Antibody preparations may be lyophilized and stored as a sterile powder, preferably under vacuum, and then reconstituted in bacteriostatic water containing, for example, benzyl alcohol preservative, or in sterile water prior to injection.

Treatment will generally involve the repeated administration of the antibody preparation via an acceptable route of administration such as intravenous injection (IV), at an effective dose. Dosages will depend upon various factors generally appreciated by those of skill in the art, including the type of cancer and the severity, grade, or stage of the cancer, the binding affinity and half life of the antibodies used, the

degree of expression of a plurality of kallikrein polypeptides, and optionally CA125, in the patient, the extent of circulating kallikrein polypeptide antigens, and optionally CA125 antigens, the desired steady-state antibody concentration level, frequency of treatment, and the influence of any chemotherapeutic agents used in combination with a treatment method of the invention.

5 Daily doses may range from about 0.1 to 100 mg/kg. Doses in the range of 10-500 mg antibodies per week may be effective and well tolerated, although even higher weekly doses may be appropriate and/or well tolerated. A determining factor in defining the appropriate dose is the amount of antibodies necessary to be therapeutically effective in a particular context. Repeated administrations may be required to achieve tumor inhibition or regression. Direct administration of antibodies specific for a plurality of kallikrein
10 polypeptides and optionally CA125 is also possible and may have advantages in certain situations.

 Patients may be evaluated for a plurality of kallikrein polypeptides and optionally CA125, preferably in serum, in order to assist in the determination of the most effective dosing regimen and related factors. The assay methods described herein, or similar assays, may be used for quantitating circulating kallikrein polypeptide and optionally CA125 levels in patients prior to treatment. Such assays may also be
15 used for monitoring throughout therapy, and may be useful to gauge therapeutic success in combination with evaluating other parameters, such as serum kallikrein polypeptides, and optionally CA125, levels.

 The invention further provides vaccines formulated to contain a plurality of kallikrein polypeptides, and optionally CA125, or fragments thereof. The use in anti-cancer therapy of tumor antigens in a vaccine for generating humoral and cell-mediated immunity is well known and, for example, has been employed in
20 prostate cancer using human PSMA and rodent PAP immunogens (Hodge et al., 1995, Int. J. Cancer 63: 231-237; Fong et al., 1997, J. Immunol. 159: 3113-3117). These methods can be practiced by employing a plurality of kallikrein polypeptides, and optionally CA125, or fragments thereof, or nucleic acids and recombinant vectors capable of expressing and appropriately presenting the kallikrein and optionally CA125, immunogens.

25 By way of example, viral gene delivery systems may be used to deliver nucleic acids encoding a plurality of kallikrein polypeptides, and optionally CA125. Various viral gene delivery systems which can be used in the practice of this aspect of the invention include, but are not limited to, vaccinia, fowlpox, canarypox, adenovirus, influenza, poliovirus, adeno-associated virus, lentivirus, and sindbus virus (Restifo, 1996, Curr. Opin. Immunol. 8: 658-663). Non-viral delivery systems may also be employed by using naked
30 DNA encoding a plurality of kallikrein polypeptides, and optionally CA125, or fragments thereof introduced into the patient (e.g., intramuscularly) to induce an anti-tumor response.

 Various *ex vivo* strategies may also be employed. One approach involves the use of cells to present kallikrein and optionally CA125 antigens to a patient's immune system. For example, autologous dendritic cells which express MHC class I and II, may be pulsed with a plurality of kallikrein polypeptides, and
35 optionally CA125, or peptides thereof that are capable of binding to MHC molecules, to thereby stimulate ovarian cancer patients' immune systems (See, for example, Tjoa et al., 1996, Prostate 28: 65-69; Murphy et al., 1996, Prostate 29: 371-380).

 Anti-idiotypic antibodies specific for a plurality of kallikrein polypeptides, and optionally CA125, can also be used in anti-cancer therapy as a vaccine for inducing an immune response to cells expressing the

polypeptides. The generation of anti-idiotypic antibodies is well known in the art and can readily be adapted to generate anti-idiotypic antibodies that mimic an epitope on a kallikrein polypeptide or CA125 (see, for example, Wagner et al., 1997, Hybridoma 16: 33-40; Foon et al., 1995, J Clin Invest 96: 334-342; Herlyn et al., 1996, Cancer Immunol Immunother 43: 65-76). Such antibodies can be used in anti-idiotypic therapy as presently practiced with other anti-idiotypic antibodies directed against tumor antigens.

Genetic immunization methods may be utilized to generate prophylactic or therapeutic humoral and cellular immune responses directed against cancer cells expressing a plurality of kallikrein polypeptides, and optionally CA125. Constructs comprising DNA encoding kallikrein and optionally CA125 polypeptides/immunogens and appropriate regulatory sequences may be injected directly into muscle or skin of an individual, such that the cells of the muscle or skin take-up the construct and express the encoded kallikrein or CA125 polypeptides/immunogens. The polypeptides/immunogens may be expressed as cell surface proteins or be secreted. Expression of the polypeptides/immunogens results in the generation of prophylactic or therapeutic humoral and cellular immunity against the cancer. Various prophylactic and therapeutic genetic immunization techniques known in the art may be used.

The invention further provides methods for inhibiting cellular activity (e.g., cell proliferation, activation, or propagation) of a cell expressing a plurality of kallikrein polypeptides, and optionally CA125. This method comprises reacting immunoconjugates of the invention (e.g., a heterogeneous or homogenous mixture) with the cell so that the kallikrein polypeptides, and optionally CA125, form complexes with the immunoconjugates. A subject with a neoplastic or preneoplastic condition can be treated when the inhibition of cellular activity results in cell death.

In another aspect, the invention provides methods for selectively inhibiting a cell expressing a plurality of kallikrein polypeptides, and optionally CA125, by reacting a combination of immunoconjugates of the invention with the cell in an amount sufficient to inhibit the cell. Amounts include those that are sufficient to kill the cell or sufficient to inhibit cell growth or proliferation.

Vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses, or from various bacterial plasmids, may be used to deliver nucleic acids encoding a plurality of kallikrein polypeptides, and optionally CA125, to a targeted organ, tissue, or cell population. Methods well known to those skilled in the art may be used to construct recombinant vectors that will express antisense nucleic acid molecules for kallikrein polypeptides and CA125. (See, for example, the techniques described in Sambrook et al (supra) and Ausubel et al (supra)).

Genes encoding a plurality of kallikrein polypeptides, and optionally CA125, can be turned off by transfecting a cell or tissue with vectors that express high levels of a desired kallikrein or CA125 polypeptide-encoding fragments. Such constructs can inundate cells with untranslatable sense or antisense sequences. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until all copies are disabled by endogenous nucleases.

Modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA or PNA, to the regulatory regions of genes encoding kallikrein polypeptides, and optionally CA125, i.e., the promoters, enhancers, and introns. Preferably, oligonucleotides are derived from the transcription initiation site, eg, between -10 and +10 regions of the leader sequence. The antisense molecules may also be designed

so that they block translation of mRNA by preventing the transcript from binding to ribosomes. Inhibition may also be achieved using "triple helix" base-pairing methodology. Triple helix pairing compromises the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Therapeutic advances using triplex DNA were reviewed by Gee J E et al (In: Huber B E and B I Carr (1994) Molecular and Immunologic Approaches, Futura Publishing Co, Mt Kisco N.Y.).

Ribozymes are enzymatic RNA molecules that catalyze the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. The invention therefore contemplates engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding a plurality of kallikrein polypeptides, and optionally CA125.

Specific ribozyme cleavage sites within any potential RNA target may initially be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once the sites are identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be determined by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Methods for introducing vectors into cells or tissues include those methods discussed herein and which are suitable for *in vivo*, *in vitro* and *ex vivo* therapy. For *ex vivo* therapy, vectors may be introduced into stem cells obtained from a patient and clonally propagated for autologous transplant into the same patient (See U.S. Pat. Nos. 5,399,493 and 5,437,994). Delivery by transfection and by liposome are well known in the art.

Kallikrein polypeptides, optionally CA125 polypeptide, and/or polynucleotides encoding the polypeptides, and fragments thereof, antibodies and/or agents identified using a method of the invention, or combinations thereof, may be used in the treatment of ovarian cancer or diseases, conditions or syndromes associated with ovarian cancer, in a subject. A combination of kallikrein polypeptides and/or polynucleotides encoding the kallikreins (e.g. kallikreins 7 and 8) and inhibitors (antisense, antibodies, or agents) of other kallikreins (e.g. kallikreins 5, 6, 10 and 11) and/or CA125 may be used in a prognostic or therapeutic method of the invention. The polypeptides, polynucleotides, and agents may be formulated into compositions for administration to subjects suffering from ovarian cancer. Therefore, the present invention also relates to a composition comprising a plurality of kallikrein polypeptides and optionally CA125, or nucleic acids encoding the polypeptides, or a fragment thereof, or an agent identified using a method of the invention, and a pharmaceutically acceptable carrier, excipient or diluent. A method for treating or preventing ovarian cancer in a subject is also provided comprising administering to a patient in need thereof, a plurality of kallikrein polypeptides and optionally CA125, or nucleic acids encoding the polypeptides, an agent identified in accordance with a method of the invention, and/or a composition of the invention.

The active substance may be administered in a convenient manner such as by injection (subcutaneous, intravenous, etc.), oral administration, inhalation, transdermal application, or rectal administration. Depending on the route of administration, the active substance may be coated in a material to

protect the substance from the action of enzymes, acids and other natural conditions that may inactivate the substance.

The compositions described herein can be prepared by per se known methods for the preparation of pharmaceutically acceptable compositions which can be administered to subjects, such that an effective
5 quantity of the active substance is combined in a mixture with a pharmaceutically acceptable vehicle. Suitable vehicles are described, for example, in Remington's Pharmaceutical Sciences (Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., USA 1985). On this basis, the compositions include, albeit not exclusively, solutions of the active substances in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered solutions with a suitable
10 pH and iso-osmotic with the physiological fluids.

The compositions are indicated as therapeutic agents either alone or in conjunction with other therapeutic agents or other forms of treatment (e.g. chemotherapy or radiotherapy). The compositions of the invention may be administered concurrently, separately, or sequentially with other therapeutic agents or therapies.

15 The following non-limiting examples are illustrative of the present invention:

Example 1

To investigate the additional discriminatory value of the kallikreins to CA125 a logistic regression model was developed. Included in the study were serum samples from 39 ovarian cancer patients and 194 non-cancer females. The age of the patients was as follows: median = 59, range 32-82. The age of the
20 controls was as follows: median = 46; range = 22-77. The model was adjusted for the following variables: $f(x) = -0.29 \text{ hK5} + 0.12 \text{ hK6} - 0.65 \text{ hK7} - 0.6 \text{ hK8} + 1.09 \text{ hK10} + 0.98 \text{ hK11} + 0.057 \text{ CA125} - 0.62$. For these data, the crude odds ratio and the 95% confidence interval were found to be 2.71 and 1.91-3.84 ($p < 0.001$). The log likelihood scores for this multivariate logistic regression model, which incorporated the combined variables for each patient was calculated. From these data, by picking different thresholds for the regression
25 function values, a ROC curve was devised which shows the added value of using kallikreins and CA125 together in a multivariate function. (AUC, 0.99; 95% CI, 0.96-1.00). (See Figure 8.) Statistically significant correlations between age and other studied variables were not observed.

Example 2

Statistically significant differences in serum kallikrein concentration was found between patient and
30 control subjects for kallikreins hK5 ($p < 0.0001$), hK7 ($p = 0.007$), hK8 ($p = 0.005$), hK10 ($p = 0.0003$) and CA125 ($p < 0.0001$) by the Mann-Whitney test. The diagnostic sensitivity (SENS) and specificity (SPEC) for each one of these markers were as follows (SENS/SPEC; both as %): 31/95 (hK5); 62/71 (hK7); 62/70 (hK10); 54/54 (hK11); 89/94 (CA125). When these data were combined in a logistic regression model, kallikreins 5 and 10 did not contribute to a great extent to the sensitivity and specificity of CA125. The area
35 under the curve of CA125 alone (93%) improved by a further 1% when adding hK6, by 2% when adding hK11, 3% when adding hK7 and 5% when adding hK8. The combination of CA125 and hK8 resulted in an AUC of 98%.

Below is a summary of each marker and its ability to separate the cases and controls.

- 34 -

- hK5: high values associated with cancer
 test+ is $hK5 > 0.10$, test- is $hK5 \leq 0.10$
 sensitivity=31%, specificity=95%,
 AUC=.62, $p(AUC)=.02$
 5 Wilcoxon rank sum test has $p < .0001$.
 Of the 233 persons analyzed, 207 have value zero for hK5 (27 cases, 180 controls).
 Possible good marker
- hK6: high values associated with cancer
 10 test+ is $hK6 > 6.3$, test- is $hK6 \leq 6.3$
 sensitivity=69%, specificity=40%,
 AUC=.50, $p(AUC)=1.00$
 Wilcoxon rank sum test has $p=.91$.
 Not a good marker
- hK7: low values associated with cancer
 15 test+ is $hK7 < 2.05$, test- is $hK7 \geq 2.05$
 sensitivity=62%, specificity=71%,
 AUC=.64, $p(AUC)=.006$
 20 Wilcoxon rank sum test has $p=.007$.
 Possible good marker
- hK8: low values associated with cancer
 25 test+ is $hK8 < 13.0$, test- is $hK8 \geq 13.0$
 sensitivity=72%, specificity=42%,
 AUC=.64, $p(AUC)=.006$
 Wilcoxon rank sum test has $p=.005$
 Possible good marker
- hK10: high values associated with cancer
 30 test+ is $hK10 > 1.42$, test- is $hK10 \leq 1.42$
 sensitivity=62%, specificity=70%,
 AUC=.68, $p(AUC)=.0004$
 Wilcoxon rank sum test has $p=.0003$.
 35 Best single kallikrein marker
- hK11: high values associated with cancer
 test+ is $hK11 > 0.14$, test- is $hK11 \leq 0.14$
 sensitivity=54%, specificity=54%,
 40 AUC=.58, $p(AUC)=.12$
 Wilcoxon rank sum test has $p=.11$.
 Not a good marker
- CA125: high values associated with cancer
 45 test+ is $Ca125 > 34$, test- is $Ca125 \leq 34$
 sensitivity=89%, specificity=94%,
 AUC=.933, $p(AUC) < .0001$
 Wilcoxon rank sum test has $p < .0001$.
 Good marker
- 50 After some further multivariate analysis of only the kallikrein markers, the combination of hK 7, 8, 10 and 11 was a preferred set. This combination was arrived at by looking at the incremental AUC as markers were combined. Below is a summary of all the models tried:
- 55 hK10 alone, AUC=.68
 hK10+hK7: AUC=.88

- 35 -

hK10+hK7+hK8: AUC=.90
 hK10+hK7+hK8+hK11: AUC=.925
 Multivariate model of hK7, hK8, hK10, hK11, call it hK7_8_10_11

- 5 hK7_8_10_11:
 Calculate $SA = 2.00 - 1.49(hK7) - .34(hK8) + 1.16(hK10) + 3.50(hK11)$
 high values associated with cancer
 test+ is $SA > -1.15$, test- is $SA \leq -1.15$
 sensitivity=87%, specificity=89%,
 10 AUC=.93, $p(AUC) \leq .0001$
 Wilcoxon rank sum test has $p < .0001$.
 Good marker

The hK marker that added the most to CA125 was also investigated.

- 15 CA125 alone, AUC=.933
 CA125+hK8: AUC=.978
 Multivariate model of Ca125, hK8, call it Ca125_hK8

- Ca125_hK8:
 20 $SC = -1.71 + .086(Ca125) - .47(hK8)$.
 high values associated with cancer
 test+ is $SC > -2.52$, test- is $SC \leq -2.52$
 sensitivity=97%, specificity=90%,
 AUC=.978, $p(AUC) \leq .0001$
 25 Wilcoxon rank sum test has $p < .0001$.
 Good marker

Below is a summary of the above analyses:

- a) The preferred kallikrein marker alone is hK10, AUC=.68
 30 b) CA125 has an AUC of .933
 c) The preferred combination of kallikrein markers increases the AUC up to .925, which is close to the CA125 AUC of .933
 d) Adding a kallikrein marker to CA125 can increase the AUC up to .978
 35 How does CA125 alone compare with the multivariate kallikrein model hK7_8_10_11?
 (based on 39 cases and 186 controls evaluated with CA125)

	Sensitivity	Specificity	misclassification
40 CA125	90%	94%	12FP, 4FN, total 16 pts misclassified
hK7_8_10_11	85%	89%	31FP, 4FN, total 35 pts misclassified
both positive	77%	100%	0FP, 9FN, total 9 pts misclassified
either positive	97%	82%	33FP, 1FN, total 34 pts misclassified

How does CA125 alone compare with the multivariate model of CA125 plus hK8?

- 45 (based on 39 cases and 186 controls evaluated with CA125)
- | | Sensitivity | Specificity | misclassification |
|-----------|-------------|-------------|---------------------------------------|
| CA125 | 90% | 94% | 12FP, 4FN, total 16 pts misclassified |
| CA125_hK8 | 95% | 91% | 17FP, 2FN, total 19 pts misclassified |

Kallikrein markers approach CA125 in terms of AUC and sensitivity, but their specificity is not as high. Adding hK8 to CA125 improves sensitivity but specificity is lower than CA125 alone.

Summary

- 5 a) The best kallikrein marker alone is hK10, area under the ROC curve (AUC) = .68.
- b) CA125 has an AUC of .933. Adding a single kallikrein marker to CA125 can get the AUC up to .978. Adding hK8 to CA125 improves sensitivity but specificity is lower compared with CA125 alone.
- 10 c) The best combination of kallikrein markers gets the AUC up to .925, which is close to the CA125 AUC of .933. Kallikrein markers approach CA125 in terms of AUC and sensitivity, but their specificity is lower.

15 The present invention is not to be limited in scope by the specific embodiments described herein, since such embodiments are intended as but single illustrations of one aspect of the invention and any functionally equivalent embodiments are within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

20 All publications, patents and patent applications referred to herein are incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety. All publications, patents and patent applications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the domains, cell lines, vectors, methodologies etc. which are reported therein

25 which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

 It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more

30 antibodies and equivalents thereof known to those skilled in the art, and so forth.

 Below full citations are set out for the references referred to in the specification.

Table 1

Kallikrein Polypeptide	Kallikrein Nucleic Acid Designation	GenBank Accession No.
Kallikrein 5	KLK5	AAD26429, AF135028, AF168768
Kallikrein 6	KLK6	AAB66483, AF013988 (CDS 174..881), AF149289 (CDS join 3567..3606, 4346..4502, 8122..8369, 9791..9927, 11805..11957) U62801 (CDS 246..980)
Kallikrein 7	KLK7	AAC37551, L33404 (CDS 16..777), AF166330 (CDS join 3237..3309, 3722..3869, 4566..4813, 5129..5265, 7362..7517) (mRNA join(1756..1785, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..8265) /product="stratum corneum chymotryptic enzyme" /note="alternatively spliced" ; mRNA join (1756..1785, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..7991) /note="alternatively spliced"; mRNA join (1821..1864, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..8265) /product="stratum corneum chymotryptic enzyme" /note="alternatively spliced"; mRNA join (1821..1864, 3179..3309, 3722..3869, 4566..4813, 5129..5265, 7362..7991) /note="alternatively spliced"
Kallikrein 8	KLK8	BAA28673, AB009849 (CDS 35..817), AF095743 (CDS join 1035..1104, 1619..1778, 1944..2206, 4304..4437, 5974..6129, mRNA 500..670, 1027..1104, 1619..1778, 1944..2206, 4304..4437, 5974..6174), AB010780 (CDS join 1..39, 418..712, 878..946), AF055982
Kallikrein 10	KLK10	AAC14266, AF055481 (CDS join 614..701, 2455..2635, 3589..3863, 4195..4328, 4793..4945, mRNA join 48..120, 605..701, 2455..2635, 3589..3863, 4195..4328, 4793..5474), NM_002776 (CDS 220..1050)
Kallikrein 11	KLK11	BAA33404, AAD47815, AB012917 (CDS 26..874), AF164623 (CDS 4224..4263, 5061..5217, 5545..5810, 6627..6763, 7158..7310) (mRNA join (2313..2398, 4189..4263, 5061..5217, 5545..5810, 6627..6763, 7158..7622)

- 38 -

Table 2

Descriptive statistics for hk5, hk6, hk7, hk8, hk10 and hk11 serum protein levels in controls and patients with ovarian cancer

	Mean	Standard Error	Median	Range	p value*
<u>hk5 (ng/ml)</u>					
Non cancer (N=194)	0.063	0.029	0.00	0.00-4.50	
Cancer (N=39)	0.48	0.18	0.00	0.00-5.70	
% Increase**	661%	--	--	--	<0.001
<u>hk6(ng/ml)</u>					
Non cancer (N=194)	6.96	0.18	6.60	1.60-15.30	
Cancer (N=39)	9.88	2.20	6.60	1.50-70.80	
% Increase**	42%	--	--	--	0.91
<u>hk7(ng/ml)</u>					
Non cancer (N=194)	2.60	0.071	2.67	0.30-6.00	
Cancer (N=39)	2.49	0.41	1.80	0.00-10.80	
% Decrease**	4%	--	33%	--	0.007
<u>hk8(ng/ml)</u>					
Non cancer (N=194)	11.74	0.27	11.70	2.40-22.20	
Cancer (N=39)	11.91	1.88	6.90	0.00-46.20	
% Decrease**	--	--	41%	--	0.005
<u>hk10(ng/ml)</u>					
Non cancer (N=194)	1.16	0.051	1.08	0.00-4.20	
Cancer (N=39)	6.51	2.46	1.59	0.27-90.0	
% Increase**	461%	--	40%	--	<0.001
<u>hk11(ng/ml)</u>					
Non cancer (N=194)	0.21	0.018	0.12	00-1.30	
Cancer (N=39)	0.79	0.21	0.18	0.00-5.52	
% Increase**	276%	--	50%		0.011

* Calculated by the Mann Whitney test

** Calculated by assuming that value in non-cancerous tissue is 100%

- 39 -

Table 3
Correlations between the studied variables in 194 non-cancer cases

variable		hK5	hK6	hK7	hK8	hK10	hK11	CA125
hK5	r_s	1.000	0.034	-0.053	0.066	0.134	0.150	0.101
	p	.	<i>0.642</i>	<i>0.462</i>	<i>0.359</i>	<i>0.062</i>	<i>0.037</i>	<i>0.172</i>
hK6	r_s	0.034	1.000	0.114	0.298	0.191	0.120	-0.160
	p	<i>0.642</i>	.	<i>0.115</i>	<i>0.000</i>	<i>0.008</i>	<i>0.097</i>	<i>0.029</i>
hK7	r_s	-0.053	0.114	1.000	0.497	0.321	0.399	0.135
	p	<i>0.462</i>	<i>0.115</i>	.	<i>0.000</i>	<i>0.000</i>	<i>0.000</i>	<i>0.066</i>
hK8	r_s	0.066	0.298	0.497	1.000	0.263	0.396	0.048
	p	<i>0.359</i>	<i>0.000</i>	<i>0.000</i>	.	<i>0.000</i>	<i>0.000</i>	<i>0.519</i>
hK10	r_s	0.134	0.191	0.321	0.263	1.000	0.176	0.035
	p	<i>0.062</i>	<i>0.008</i>	<i>0.000</i>	<i>0.000</i>	.	<i>0.014</i>	<i>0.638</i>
hK11	r_s	0.150	0.120	0.399	0.396	0.176	1.000	0.125
	p	<i>0.037</i>	<i>0.097</i>	<i>0.000</i>	<i>0.000</i>	<i>0.014</i>	.	<i>0.090</i>
CA125	r_s	0.101	-0.160	0.135	0.048	0.035	0.125	1.000
	p	<i>0.172</i>	<i>0.029</i>	<i>0.066</i>	<i>0.519</i>	<i>0.638</i>	<i>0.090</i>	.

Table 4
Correlations between the studied variables in 39 ovarian cancer cases

variable		hK5	hK6	hK7	hK8	hK10	hK11	CA125
hK5	r_s	1.000	0.475	0.553	0.554	0.618	0.584	0.507
	p	.	0.002	0.000	0.000	0.000	0.000	0.001
hK6	r_s	0.475	1.000	0.327	0.513	0.470	0.661	0.530
	p	0.002	.	0.042	0.001	0.003	0.000	0.001
hK7	r_s	0.553	0.327	1.000	0.695	0.690	0.748	0.262
	p	0.000	0.042	.	0.000	0.000	0.000	0.107
hK8	r_s	0.554	0.513	0.695	1.000	0.602	0.783	0.443
	p	0.000	0.001	0.000	.	0.000	0.000	0.005
hK10	r_s	0.618	0.470	0.690	0.602	1.000	0.706	0.548
	p	0.000	0.003	0.000	0.000	.	0.000	0.000
hK11	r_s	0.584	0.661	0.748	0.783	0.706	1.000	0.556
	p	0.000	0.000	0.000	0.000	0.000	.	0.000
CA125	r_s	0.507	0.530	0.262	0.443	0.548	0.556	1.000
	p	0.001	0.001	0.107	0.005	0.000	0.000	.

Table 5

		CANCER				CONTROL				Wilcoxon										
	n	mean	median	sd	min	max	n	mean	median	sd	min	max	p(mean)	p(logistic)	AUC (se)	p(AUC)	OUTPOINT	SENSITIVITY	SPECIFICITY	
	39	0.48	0.00	1.14	0.00	5.70	194	0.06	0.00	0.40	0.00	4.50	0.03	<0.001	0.005	0.62 (.05)	0.0200	>0.10	31%	95% HK5
HK5	39	9.88	6.60	13.77	1.50	70.80	194	6.97	6.60	2.60	1.60	15.30	0.20	0.9100	0.0570	0.50 (.05)	1.0000	>6.30	69%	40% HK6
HK6	39	2.48	1.80	2.54	0.00	10.80	194	2.60	2.67	0.99	0.30	6.00	0.79	0.0070	0.6500	0.64 (.05)	0.0000	>6.30	62%	71% HK7
HK7	39	11.91	6.90	11.80	0.00	46.20	194	11.74	11.70	3.60	2.40	22.20	0.93	0.0050	0.8700	0.64 (.05)	0.0050	<13.0	42%	42% HK8
HK8	39	0.52	1.59	15.40	0.27	90.00	194	1.17	1.08	0.72	0.00	4.20	0.04	0.0003	0.0002	0.68 (.05)	0.0004	>1.42	62%	70% HK10
HK10	39	0.80	0.18	1.36	0.00	5.52	194	0.21	0.12	0.26	0.00	1.30	0.01	0.1100	0.0003	0.58 (.05)	0.1200	>0.14	54%	54% HK11
HK11	39	0.49	2.77	9.71	20	4494	186	24	21	9	0	84	0.00	<0.001	<0.001	0.933 (.026)	<0.001	>34	90%	94% C4125
C4125	39	4.51	4.51	15.81	4.74	92.26	194	3.84	3.68	2.07	<0.05	1.07	0.002	<0.001	<0.01	0.925 (.030)	<0.001	>1.32	85%	60% HK7 & 10 11
HK7 & 10 11	39	48.64	15.02	82.86	-5.86	382.85	186	-5.07	-4.92	1.91	-9.49	0.50	0.0002	<0.001	<0.01	0.978 (.017)	<0.001	>2.52	97%	90% C4125, HK8

We Claim:

1. A method for detecting a plurality of kallikrein markers associated with ovarian cancer in a patient comprising:
 - 5 (a) obtaining a sample from a patient;
 - (b) detecting in the sample a plurality of kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
 - (c) comparing the detected amounts with amounts detected for a standard.
- 10 2. A method for diagnosing and monitoring ovarian cancer in a subject comprising detecting in a sample from the subject a plurality of kallikrein markers, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein.
3. A method as claimed in claim 1 or 2 wherein the plurality of kallikrein markers are detected using
 - 15 antibodies that bind to each of the plurality of kallikrein markers or parts thereof
4. A method as claimed in claim 1, 2, 3 which further comprises detecting CA125.
5. A method of detecting ovarian cancer in a patient, the method comprising comparing:
 - 20 (a) levels of a plurality of kallikrein markers, and optionally CA125, in a sample from the patient, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and
 - (b) normal levels of expression of the plurality of kallikrein markers, and optionally CA125, in a control sample, wherein a significant difference in levels of kallikrein markers and optionally CA125, relative to the corresponding normal levels, is indicative of ovarian cancer.
- 25 6. A method for monitoring the progression of ovarian cancer in a patient, the method comprising: (a) detecting in a sample from the patient at a first time point, a plurality of kallikrein markers, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; (b) repeating step (a) at a subsequent point in time; and (c) comparing levels detected in steps (a) and (b), and thereby monitoring the
 - 30 progression of ovarian cancer.
7. A method for determining in a patient whether ovarian cancer has metastasized or is likely to metastasize in the future, the method comprising comparing (a) levels of a plurality of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are
 - 35 selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (b) normal levels or non-metastatic levels of the kallikrein markers and optionally CA125, in a control sample wherein a significant difference between the levels of expression in the patient sample and the normal levels or non-metastatic levels is an indication that the ovarian cancer has metastasized.
8. A method for assessing the aggressiveness or indolence of ovarian cancer comprising comparing:

- (a) levels of expression of a plurality of kallikrein markers, and optionally CA125, in a patient sample, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (b) normal levels of expression of the plurality of markers and optionally CA125, in a control sample, wherein a significant difference between the levels in the patient sample and normal levels is an indication that the cancer is aggressive or indolent.
- 5
9. A method for diagnosing and monitoring ovarian cancer in a sample from a subject comprising isolating nucleic acids from the sample; and detecting in the sample polynucleotides encoding a plurality of kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein.
- 10
10. A method as claimed in claim 9 wherein significant differences in the levels of the polynucleotides in the sample compared to a control is indicative of disease, disease stage, and/or prognosis.
11. A method for determining the presence or absence of ovarian cancer in a subject comprising: (a) contacting a sample obtained from the subject with oligonucleotides that hybridize to polynucleotides encoding kallikrein markers, and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (b) detecting in the sample a level of nucleic acids in the sample that hybridize to the polynucleotides relative to a predetermined cut-off value, and therefrom determining the presence or absence of ovarian cancer in the subject.
- 15
- 20
12. A method as claimed in claim 11, wherein the nucleic acids are mRNA and the levels of nucleic acids are detected by polymerase chain reaction.
13. A method as claimed in claim 11 wherein the nucleic acids are mRNA and the amounts of mRNA are detected using a hybridization technique, employing oligonucleotide probes that hybridize to kallikrein markers, and optionally CA125.
- 25
14. A method for assessing the potential efficacy of a test agent for inhibiting ovarian cancer in a patient, the method comprising comparing: (a) levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a first sample obtained from a patient and exposed to the test agent, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and (b) levels of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a second sample obtained from the patient, wherein the sample is not exposed to the test agent, wherein a significant difference in the levels of expression of the plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the first sample, relative to the second sample, is an indication that the test agent is potentially efficacious for inhibiting ovarian cancer in the patient.
- 30
- 35
15. A method of claim 14 wherein the first and second samples are portions of a single sample obtained from the patient.
16. A method of claim 14 wherein the first and second samples are portions of pooled samples obtained

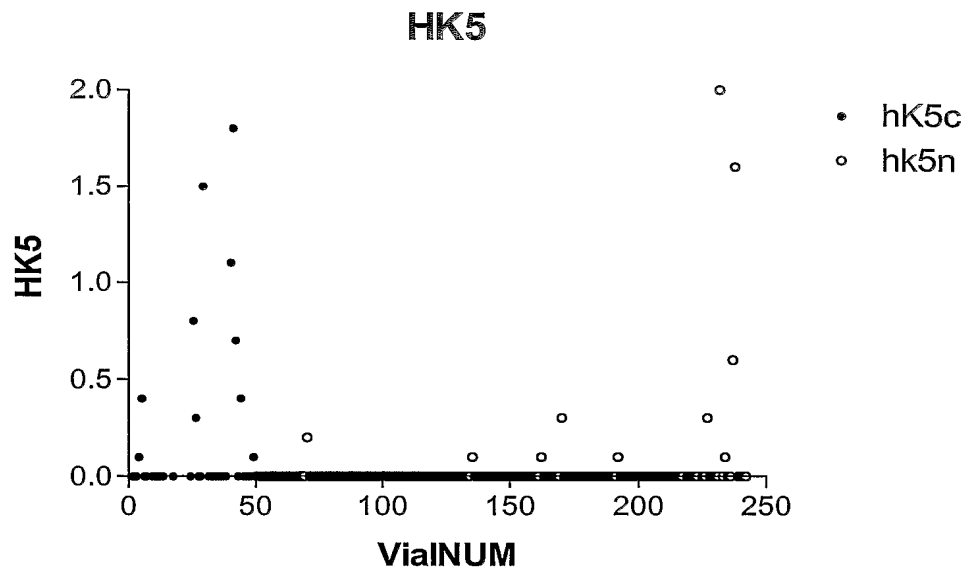
from the patient.

17. A method of assessing the efficacy of a therapy for inhibiting ovarian cancer in a patient, the method comprising comparing: (a) levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a first sample obtained from the patient, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and (b) levels of the kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in a second sample obtained from the patient following therapy, wherein a significant difference in the levels of expression of the kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the second sample, relative to the first sample, is an indication that the therapy is efficacious for inhibiting ovarian cancer in the patient.
18. A method of selecting an agent for inhibiting ovarian cancer in a patient the method comprising (a) obtaining a sample comprising cancer cells from the patient; (b) separately exposing aliquots of the sample in the presence of a plurality of test agents; (c) comparing levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (d) selecting one of the test agents which alters the levels of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot containing that test agent, relative to other test agents.
19. A method of inhibiting ovarian cancer in a patient, the method comprising (a) obtaining a sample comprising cancer cells from the patient; (b) separately maintaining aliquots of the sample in the presence of a plurality of test agents; (c) comparing levels of a plurality of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; and (d) administering to the patient at least one of the test agents which alters the levels of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot containing that test agent, relative to other test agents.
20. A method of assessing the ovarian cell carcinogenic potential of a test compound, the method comprising: (a) maintaining separate aliquots of ovarian cells in the presence and absence of the test compound; and (b) comparing expression of a plurality of markers, optionally CA125, and/or polynucleotides encoding same, in each of the aliquots, wherein the markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11, and wherein a significant difference in levels of kallikrein markers, optionally CA125, and/or polynucleotides encoding same, in the aliquot maintained in the presence of the test compound, relative to the aliquot maintained in the absence of the test compound, is an indication that the test compound possesses ovarian cell carcinogenic potential.
21. A method of inhibiting ovarian cancer in a patient at risk for developing ovarian cancer, the method comprising inhibiting expression of genes encoding kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5,

- kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.
22. A method of any preceding claim wherein the plurality comprises at least three of the markers.
23. A method of any preceding claim wherein the plurality comprises at least five of the markers.
24. A method of any preceding claim wherein the plurality of kallikrein markers is selected from the
5 group consisting of kallikrein 5, kallikrein 7, and kallikrein 8; kallikrein 5, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 5, kallikrein 7, kallikrein 8, and kallikrein 10; kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; kallikrein 5, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11; or kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.
- 10 25. A method of any proceeding claims wherein the kallikrein markers are kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11.
26. A method of any preceding claim wherein the patient sample comprises serum obtained from the patient.
27. A kit for carrying out a method as claimed in any preceding claim.
- 15 28. A kit for assessing whether a patient is afflicted with ovarian cancer, the kit comprising reagents that specifically bind with a plurality of kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.
- 20 29. A kit for assessing the suitability of each of a plurality of agents for inhibiting ovarian cancer in a patient, the kit comprising: (a) the plurality of agents; and (b) reagents for detecting a plurality of kallikrein markers and optionally CA125, wherein the kallikrein markers comprise or are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10, and kallikrein 11.
- 25 30. A kit as claimed in claim 28 or 29 wherein the reagents are antibodies that specifically bind with protein or protein fragments corresponding to kallikrein markers and optionally CA125.

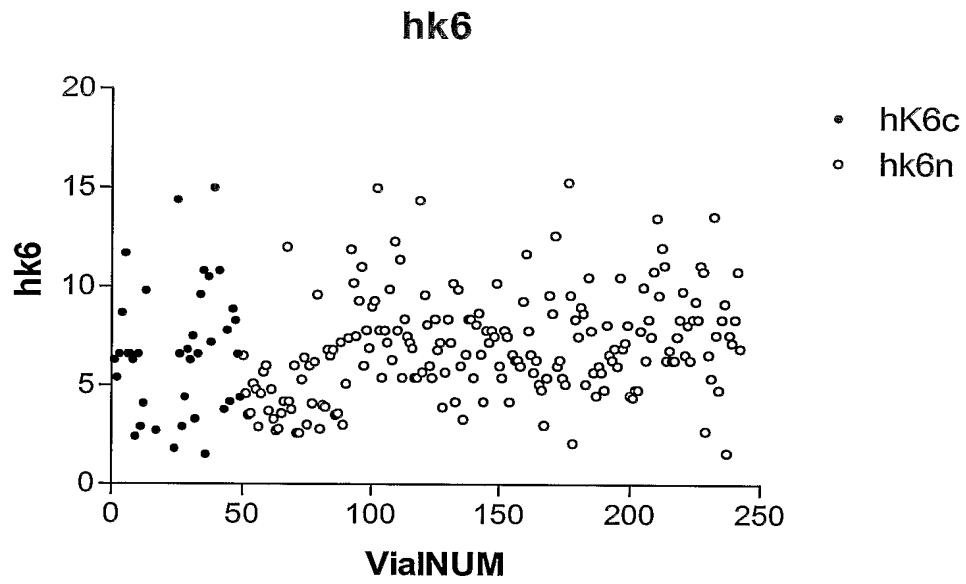
1/8

Figure 1



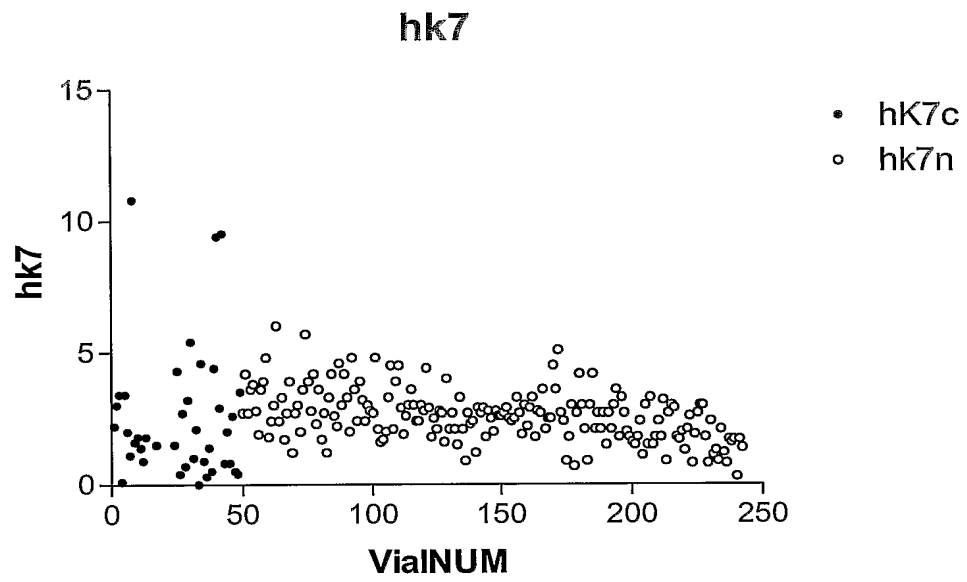
2/8

Figure 2



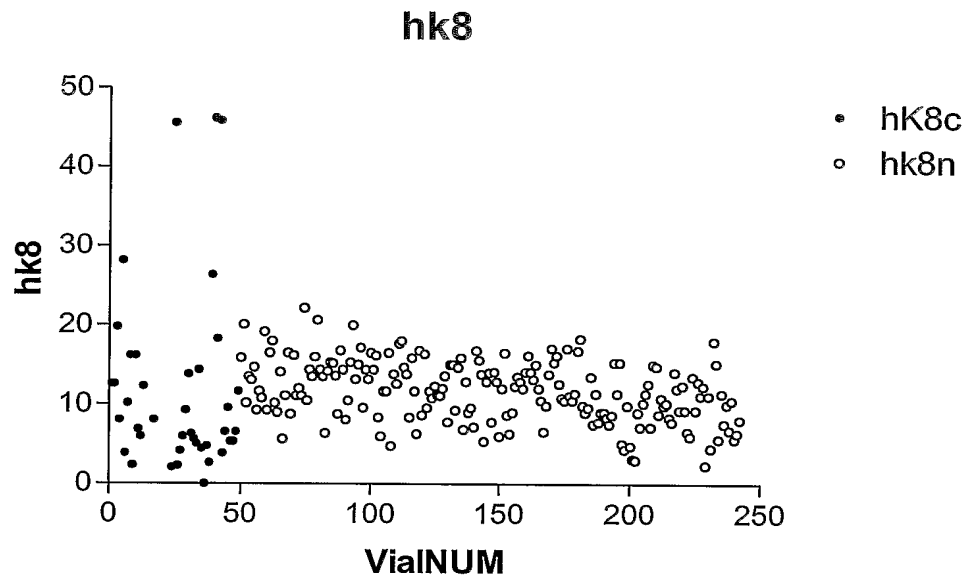
3/8

Figure 3



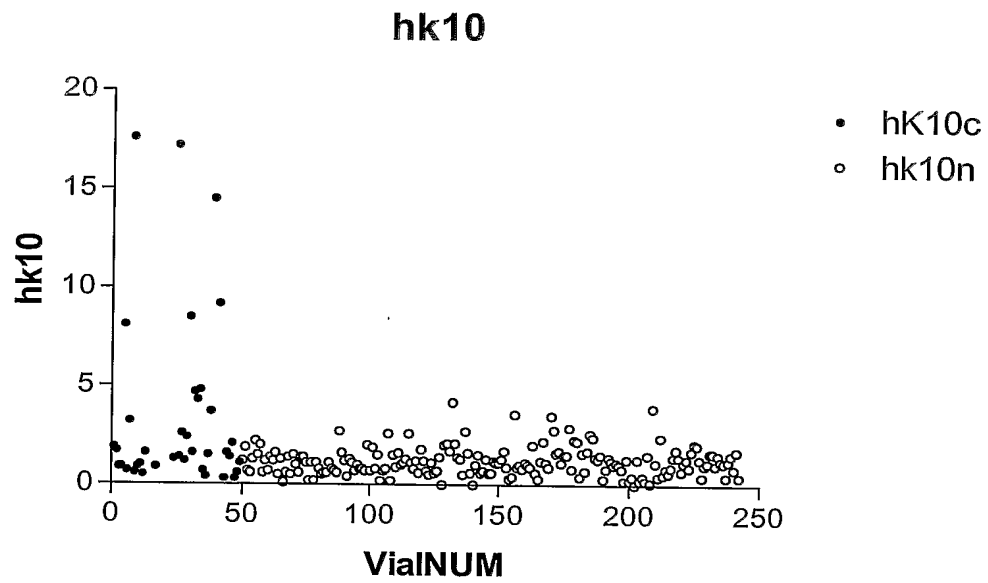
4/8

Figure 4



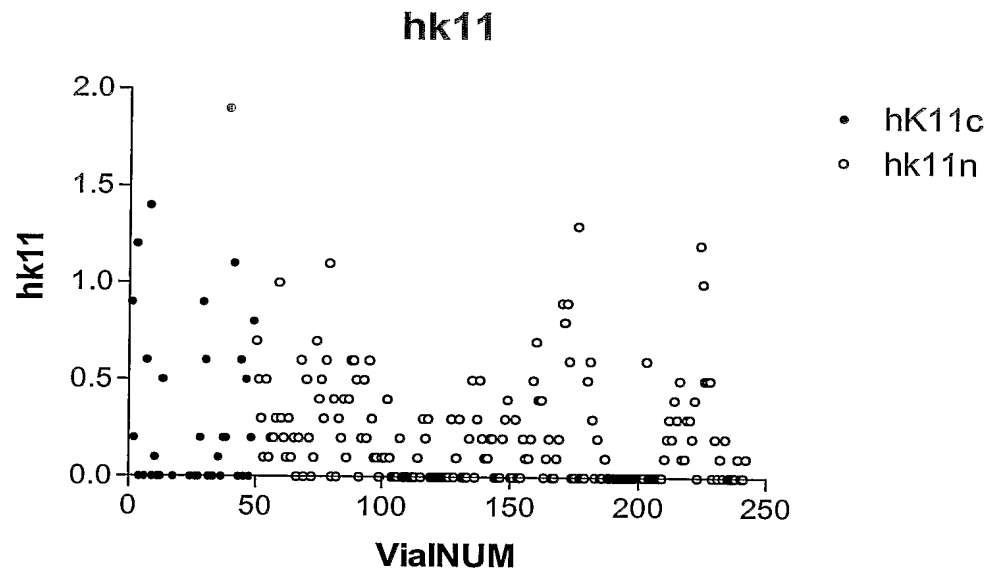
5/8

Figure 5



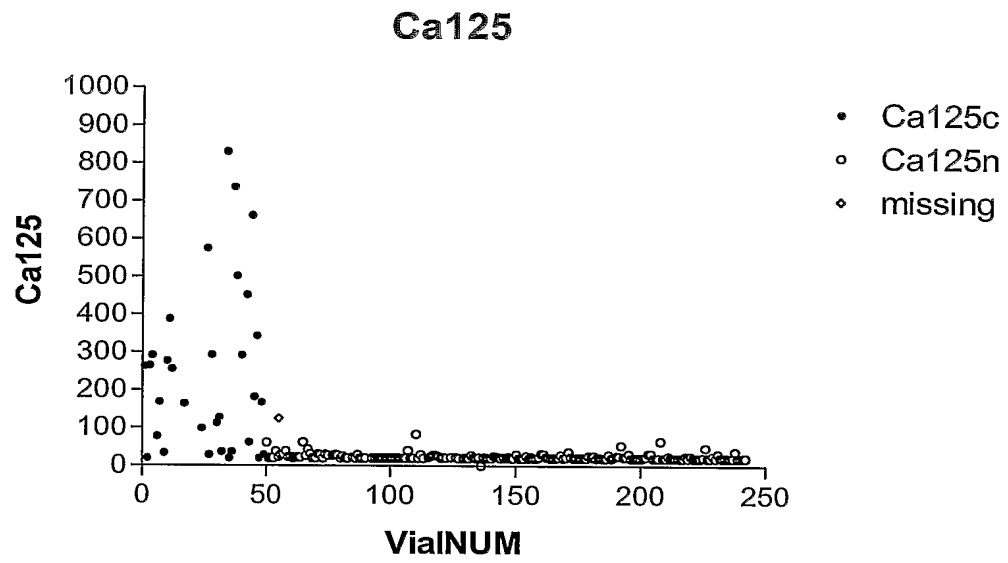
6/8

Figure 6



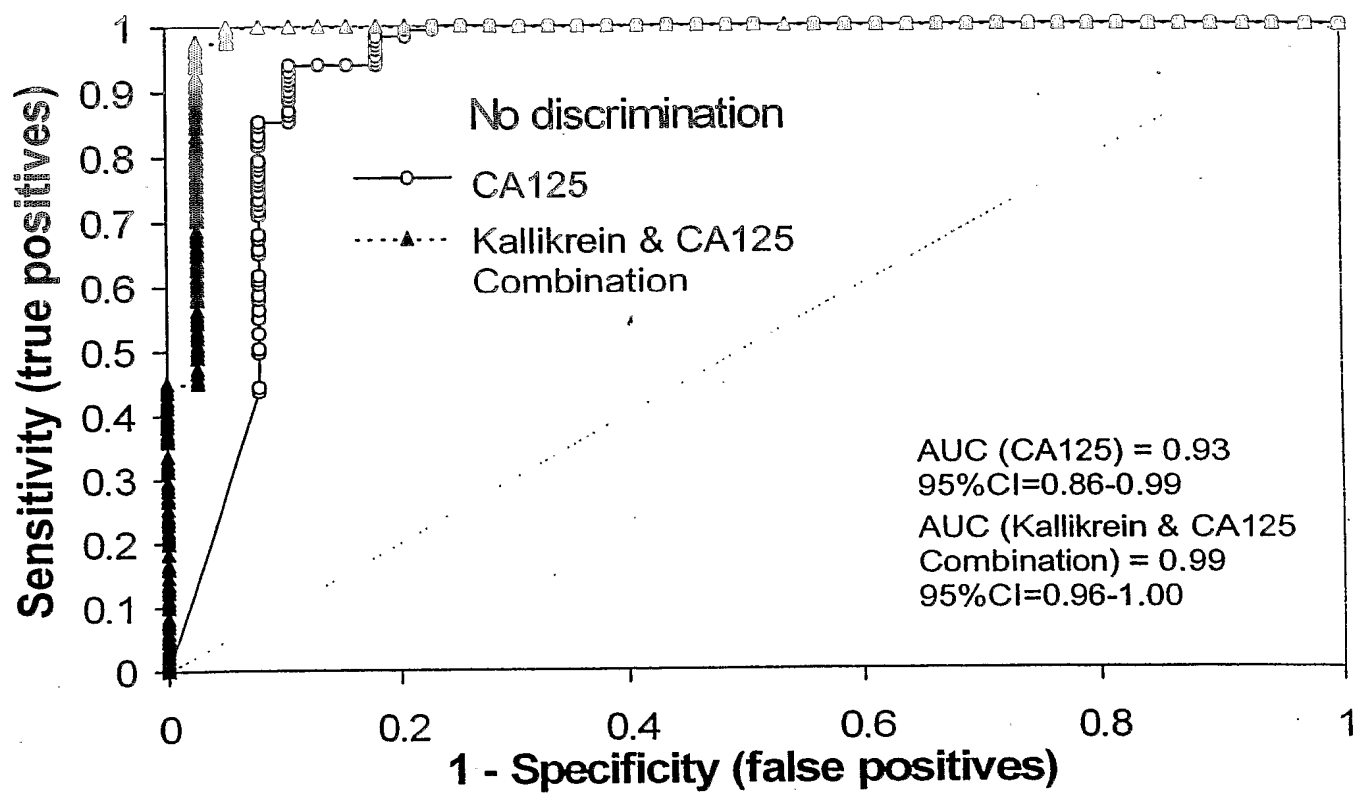
7/8

Figure 7



8/8

Figure 8



Sequence Listing

SEQ ID NO. 1

CA125 amino acid

```

5      1 mlkpsglpgs ssptrslmtg srstkatpem dsgrtgatls pktstgaivv tehtlpftsp
      61 dktlasptss vvgrrttqslg vmssalpest srgmthseqr tpslspqvn gtpsrnypat
      121 smvsglsspr trtsstegnf tkeastytlt vettsgpvte kytvptetst tegdstetpw
      181 dtryipvkit spmktfadst askenapvsm tpaettvtds htpgrtnpsf gtlyssfldl
      241 spkgtpnsrg etslelilst tgyfsspep gsaghsrist saplssasv ldnkisetst
10     301 fsgqsltspl spgvpearas tmpnsaifp mtlснаetsa ervrstissl gtpsisstkqt
      361 aetiltfhaf aetmdipsth iaktlasewl gspgtlpgts tsalttttsp tllvseentt
      421 hhstsgkete gtlntsmtpl etsapegeese mtatlvpitg fttldskirs psqvssshpt
      481 relrttgsts grqssstaah gssdilratt sstskasswt sestaqqfse pghtqgwets
      541 psmkterppa stsvaapitt svpsvsgft tlktsstkgi wleetsadtl igestagptt
15     601 hqfavptgis mtggsstrgs qgtthlltra tassetsadl tlatngvpvs vspavsktaa
      661 gssppggtkp sytmvssvip etsslgssaf regtsgltp lntrhpfssp epdsaghtki
      721 stsipllssa svledkvsat stfshhkats sittgtpeis tktkpssavl ssmtlsnaat
      781 spervrnats plthpspsge etagsvltls tsaettdspn ihptgtltse ssespstls
      841 psvsgvkttf ssstpsthlf tsgeeteets npsvsqpets vsrvrttlas tsvptpvfpt
20     901 mdtwptrsaq fssshlvsel ratsstsvtn stgsalpkis hltgtatmsq tnrdtfnDSA
      961 apqsttwpet sprfktglps attvtstsat slsatvmvsk ftspatssme atsirepstt
      1021 ilttettngp gsmavastni pigkgyiteg rldtshlpig ttassetsmd ftmakesvsm
      1081 svspsqsmda agsstpgrts qfvdtfsddv yhltsreiti prdgtssalt pgmtathpps
      1141 pdpgsarstw lgilssspss ptpkvtsmst fstqrvttsm imdtvetsrw nmpnlpstts
25     1201 ltpsnipstg aigkstlvpl dtpspatsle asegglptls typestntps ihlgahasae
      1261 spstikltma svvkpgsytp ltfpsiehti hvstarmays sgsspemtap getntgstwd
      1321 pttytittdp kdtssaqvst phsvrtlrrt enhpktesat paaysgspki ssspnltspa
      1381 tkawtitdtt ehstqlhytk laekssgfet qsagppsvsvv iptspitgss tleltsdvp
      1441 eplvlapseq ttitlpmatw lstslteema stdldissps spmstfaifp pmstpsshels
30     1501 kseadtair ntdsttldqh lgirslgrtg dlttvpitpl tttwtsvieh stqaqdtlsa
      1561 tmspthvtqs lkdqtsipas aspshlvey pelgtqgrss seattfwkps tdtlsreiet
      1621 gptniqstpp mdntttgsss sgvtlgiah lgtsspaet stnmalerrs statvsmagt
      1681 mglvtsapg rsisqslgrv ssvlseste gvtsskgs prlntqgnta lsslepsy
      1741 egsgmstsis ltsspttpdv efiggstfw kevttvmtsd iskssartes ssatlmstal
35     1801 gstentgkek lrtasmdlps ptpsmevtpw isltlnapn ttdsldlshg vhtssagtl
      1861 tdrslntgvt rasrlengsd tsskslsmgn sthtsmtde ksevsstihp rpetsappgae
      1921 tltstpgnr aistltpfss ipveevistg itsgpdinsa pmthspitpp tivwtstgti
      1981 eqstqplhav ssekvsvtq stpyvnsvav saspthensv ssgsstsspy ssasleslds
      2041 tisrrnaits wlwdlttslp tttwpstsls ealssghsgv snpsstttef plfsaastsa
40     2101 akqrnpetst hgpqntaast lntdassvtg lsetpvgsi ssevplpmi tssrdvsglt
      2161 sestansplg tassagtklt rtislptses lvsfrmnkdp wtvsiplgsh pttntetsip
      2221 vnsagppgls tvasdvidtp sdgaesiptv sfspspdtev ttishfpekt thsfrtissl
      2281 theltsrvtp ipgdwmssam stkptgasps itlgerrrit saapttspiv ltasfttetst
      2341 vsldnettkv tsdildarkt nelpsdssss sdlintsias stmdvktkas isptsisgmt
45     2401 asspslfs drpqvptst etntatpsv ssntysldgg snvggtpstl ppftithpve
      2461 tssallawsr pvrftstmvs tdtasgenpt ssnsvtvsv apgtwasvgs ttdlpamgfl
      2521 ktspageahs llastiepat aftphlsaav vtgssatsea slttteska ihsspqtptt
      2581 ptsganwets atpesllvvt etsdtlttsk ilvtdtilfs tvstppsckfp stgtlsgasf
      2641 ptllpdtpai pltateptss latsfdstpl vtiasdslt vpettltmse tsngdalvlk
50     2701 tvsnpdrip gitigqvtes plhpsstsp kivaprrnty egstvalst lpagttgslv
      2761 fsgssenset talvdssagl erasvmlptt gsqgmassgg irsgststg tktfsslplt
      2821 mnpgevtams eittnrltat gstatkgipv kptsaesgll tpvsassps kafaslttap
      2881 pstwgipqst ltfefsevp ldtksaslpt pgqslntipd sdastasssl skspeknpra
      2941 rmmntstkais assfgtgft etpegsasps magheprvpt sgtgdpryas esmsypdpsk
55     3001 assamtstsl asklttlfst ggaarsgss spislsteke tsflsptast srktslflgp
      3061 smarqpnilv hlqtsaltls ptstlnmsqe eppeltssqt iaeeegttae tqtlftftpe
      3121 tpsllpvss pteptarrks spetwassis vpaktslvet tdgtlvttik mssqaaqgns
      3181 twpapaeetg tspagtspps pevsttlkim sskepsispe irstvrnspw ktpettvpme
      3241 ttvepvtlqs talgsgtsi shlptgttsp tksptenmla tervslspsp peawtnlysg

```


	3301	tpggtrqsla	tmssvslesp	tarsitgtgq	qsspelvskt	tgmefsmwhg	stgggtgdth
	3361	vslstssnil	edpvtspns	ssltdkshhk	tetwvsttai	pstvlnnkim	aaeqqtsrsv
	3421	deaysstssw	sdqtsqsdit	lgaspdvnt	lyitstaqtt	slvslpsgdq	gitsltnpsg
	3481	gktssassvt	spsigletlr	anvsavksdi	aptaghlsqt	sspaevsild	vttaptpgis
5	3541	ttittmgtns	istttpnpev	gmstmdstpa	terrttsteh	pstwsstaas	dswtvtmdts
	3601	nlkvarspgt	istmhttsfl	assteldsms	tphgritvig	tslvtppsda	savktetsts
	3661	ertlpsdtt	astpistfsr	vqrmsisvpd	ilstswtpss	teaedvpvsm	vstdhastkt
	3721	dpntplstfl	fdslstldwd	tgrslssata	ttsapqgatt	pgeltletmi	spatsqlpfs
	3781	ighitsavtp	aamarssgvt	fsrpdptskk	aeqtstqlpt	ttsahpgqvp	rsaattldvi
10	3841	phtaktptdat	fqrqqqtalt	tearatsdsw	nekekstpsa	pwitemmns	sedlikevts
	3901	sssvlkdpay	aghklgiwdd	fipkfgkaah	mrelpllspp	qdkeaihpst	ntvettgwt
	3961	ssehashsti	pahsasklt	spvttsttre	qaivsmsttt	wpestrarte	pnsflltielr
	4021	dvspymdtss	ttqtsiissp	gstaitkgpr	teitsskris	ssflaqsmrs	sdpseaitr
	4081	lsnfpamtes	ggmilamqts	pggatslsap	tldtsatasw	tgtplattqr	ftysekttlf
15	4141	skgpedtsqp	sppsveetss	ssslvpihat	tspsnillts	qghspstpp	vtsvflsets
	4201	glgktttdmsr	islepgtslp	pnlsstagea	lstyeasrdt	kaihhsadta	vtnmeatss
	4261	yspipghktk	skatsplvts	hingditsst	svfgssette	ietvsvnqg	lqerstsqva
	4321	ssatetstvi	thvssgdatt	hvtktqatfs	sgtsissphq	fitstntftd	vstnpstsl
	4381	mtessgvtit	tqtgptgaat	qgpylltdst	mpyltetpla	vtpdfmqsek	ttliskgpkd
20	4441	vtwtsppsva	etsypssltp	flvttippat	stlqqghtss	pvsatsvlts	glvkttdmln
	4501	tsmepvtmsp	qnltnpsnei	latlaattdd	etihpsinka	vtnmgtassa	hvlhstlpvs
	4561	sepstatspm	vpasmgdal	asisipgset	tdiegeptss	ltagrkenst	lqemnsttes
	4621	niilsnsvsg	aiteatkmev	psfdatfipt	paqstkfpdi	fsvassrlsn	sppmtisthm
	4681	tttqtgssga	tskipalldt	stletsagtp	svvtegfahs	kittamndv	kdvsgttnppf
25	4741	qdeasspsq	apvlvtlpls	svaftpqwhs	tsspvsmssv	ltsslvktag	kvdtsletvt
	4801	sspqmsntl	ddisvtsaat	tdietthpsi	ntvvtnvgtt	gsafeshstv	saypepskvt
	4861	spnvttstme	dttsirsipk	ssktrrtete	ttssltpklr	etsisqeits	stetstvpky
	4921	eltgattevs	rtldvtsst	sfpqpdqstv	sldistetnt	rlstspimte	saeitittqt
	4981	gphgatsqdt	ftmdpsnttp	qagihsamth	gfsqldvttl	msripqdvsw	tsspsvdkts
30	5041	spssflsspa	mttpslisst	lpedklsspm	tslltsglvk	itdilrtrle	pvtsslpnfs
	5101	stsdakilats	kdkskdtkeif	psinteetnv	kannsgghesh	spaladsetp	kattqmvitt
	5161	itvgdpapsts	mpvhgssett	nikreptyfl	tprlretsts	qessfpdtds	fllskvptgt
	5221	itevsstgvn	ssskistpdh	dkstvppdtf	tgeiprvfts	siktksaemt	ittqasppes
	5281	ashstlpldt	sttlsgqgth	stvtqgfpys	evttlmgm	gnvswmttpp	veatssvssl
35	5341	msspamtsp	pvsstspqsi	pssplpvtal	ptsvlvtttd	vlgttspesv	tssppnlssi
	5401	therpatykd	tahteaaamhh	stntavtnvg	tsqsgghksqs	svladsetsk	atplmsttst
	5461	lgdtsvstst	pnisqtnqiq	teptaslspr	lresstsekt	ssttetntaf	syvptgaitq
	5521	asrteissr	tsisdldrpt	iapdistgmi	trlftspimt	ksaemtvtq	tttpgatsqg
	5581	ilpwtsttl	fqqgthstvs	qgfphseitt	lrsrtpgdvs	wmttppveet	ssgfsmlmps
40	5641	mtspspvss	spesipsspl	pvtalltsvl	vtttnvlgtt	spetvtsspp	nlssptqerl
	5701	ttkykdahte	amhasmhtnt	avanvgtsis	ghesqssvpa	dshtskatsp	mgitfamgdt
	5761	svststpaff	etriqtests	slipglrdtr	tseeintvte	tstvlsevpt	ttttevsrte
	5821	vitssrttis	gpdhskmspy	istetitrls	tfpftvgste	maitnqtgpi	gtisqatltil
	5881	dtssstasweg	thspvtgrfp	hseetttsmr	stkgvswqsp	psveetssps	spvplpait
45	5941	hsslysavsg	ssptsalpvt	slltsgrrrkt	idmldthsel	vtsslpssas	fsgeiltsea
	6001	stntetihfs	entaetnmgt	tnsmhklhss	vsihsqpsgh	tppkvtgsmm	edaivststp
	6061	gspetknvdr	dstspltpel	kedstalvmn	sttesntvfs	svsladaatev	sraevtyydp
	6121	tfmpasaqst	kspdispeas	sshnsnpplt	isthktiatq	tgpsgvtslg	qltldtstia
	6181	tsagtpart	qdfvdsetts	vmnndlnvdl	ktspfsaeaa	nlssqapll	vtspspvts
50	6241	tlqehstssl	vsvtsvptpt	lakitdmtdn	lepvttrspqn	lrntlatsea	tttdthmhrs
	6301	intamanvgt	tsspnefyft	vspdsdpyka	tsavvitats	gdsivstsm	rssamkkies
	6361	ettfslifrl	retstsqkig	sssdstsvfd	kaftaattev	srteltsssr	tsiqgtekpt
	6421	mtpdtstrsv	tmlstfaglt	kseertiatq	tgphratsqg	tltdwtstitt	sqagthsam
	6481	hgfsqldlst	ltsrvpeyis	gtsppsvekt	ssssllslp	aitsppvpt	tlpesrpssp
55	6541	vhltslptsg	lvktttdmlas	vaslppnlgs	tshkipttse	dikdtekmy	stniavtnvg
	6601	tttsekesys	svpayseppk	vtspmvtsfn	irtdivstsm	pgsseitrie	mestfsvahg
	6661	lkgtsstsqd	ivsteksavl	hklttgatet	srtevassrr	tsipgpdhst	espdistevi
	6721	pslpislgit	essnmliitr	tgpplgstsq	gtftldtptt	ssragthsm	tqefphsemt
	6781	tvnmkdpeil	swtippsiek	tsfssslmps	pamtsppvss	tlpktihhttp	spmtslilts
60	6841	lvmttdtlgt	spepttsspp	nlstshvil	ttddedttai	amhpststaa	tnvettcsgh
	6901	gsqssvltlds	ektkatapmd	ttstmghtt	stmsvssset	tkikrestys	ltpglretsi

5 6961 sqnasfsttdt sivilsevpptg ttaevsrtev tssgrtsipg psqstvlpei strtmtrlfa
 7021 sptmtesaem tiptqtgpgsg stsqdtltld tsttksgakt hstltqrfph semttlmsrg
 7081 pgdmwqssp slenpsslps llslpattsp ppisstlpvt isssplpvts lltsspvttt
 7141 dmlhtspelv tssppklshst sderlittgkd ttnteavhps tntaasnvei psfghesps
 7201 aladsetska tspmfitstq edttvaistp hfletsriqk esisslsplk retgssvets
 7261 saietsavls evsigattei srtevtsssr tsisgsaest mlpeisttrk iikfptspil
 7321 aessentikt qtsppgstse stftldtstt pslvithstm tqrlphseit tlvsrgagdv
 7381 prpsslpvee tsppssqlsl samispspvs stlpasshss sasvtspltp gqvkttevd
 7441 asaepetssp pslsstsvei latsevttdt ekihpfpnta vtkvgtsessg hespsvlpd
 10 7501 settkatsam gtisimgdts vstltplasn trkiqsepas slttrlrrets tseetslate
 7561 antvlskvst gattevsrte aisfsrtsms gpeqstmsqd isigtipris assvltesak
 7621 mtittqtgpps estlestlnl ntattpswve thsiviqgfp hpemttsmgr gpggvswpsp
 7681 pfvketspps splslpavts phpvsttfla hippslpvt slttsgpatt tdilgtstep
 7741 gtssssslst tsherlittyk dtahteavhp stntggtnva ttssgyksqs svladsspmc
 15 7801 ttstmgdtsv ltstpaflet rriqtelass ltpglressg segtssgatk stvlskvptg
 7861 atteiskdv tsipgpaqst ispdistrtrv swfstspvmt esaeitmnh tsplgattqg
 7921 tstlatsstt sltmthstis qgfshsqmst lmrrgpedvs wmsppllekt rpsfslmssp
 7981 attspspvss tlpesisssp lpvtslltsg lakttdmlhk ssepvtnsa nlsstsvail
 8041 atsevttdte kthpssnrtv tdvgtsssg estsfvlads qtskvtspmv itstmedtsv
 20 8101 ststpgffet sriqteptss ltlglrktss segtslatem stvlsgvptg ataevsrtev
 8161 tsssrtsisg faqltvspet stetitrlpt ssimtesaem miktqtdppg stpesthtvd
 8221 isttpnwvet hstvtqrfs hstvtqlvrs pgdmlwpsqs sveetssass llslpattsp
 8281 spvsstlved fpsaslpvts lltpglvitt drmgisrepg tsstsnlsst sherlittled
 8341 tvdtedmpps thtavnrtv sigshesqss vlsdsetpka tspmgttytm getsvsists
 25 8401 dffetsriqi eptssltsgl retssseris sategstvlsv evpsgattev srtevisrg
 8461 tsmgpdqft ispdisteai trlstspimt esaesaitie tgspgatseg tlldtsttt
 8521 fwsghstas pgfshsemmt lmsrtpgdvp wpslpveea ssvssslssp amtstsfssa
 8581 lpesissph pvtalltlgp vktdmlrtv sepetsppn lsstsaaila tsevtkdrek
 8641 ihpssntpvv nvgtviykh lpsssvladv spssvladv tskptspmat tstpafpetm
 30 8701 mtqptsslts glreistsqe tssatersas lsgmptgatt kvartealsl grtstpgpaq
 8761 stispeiste titristplt ttgsaemt it pktghsgass qgtftldtss raswpghsa
 8821 athrsphsgm ttpmsrgped vswpsrpsve ktsppsslvsv lsavtspsspl ystpsesshs
 8881 splrvtslft pvmmkttldml dtslepvtts ppsmnitsde slatskatme teaiqlsent
 35 8941 avtqmgtsa rgefysyypg lpepskvts vvtssstikdi vsttipasse itriemests
 9001 tltpptrets tsqeihsatk pstvpykalt satiedsm tq vmsssrpgsp dqstmsqdis
 9061 sevitrlists pikaestemt ittqtgspga tsrgtltldt sttfmsgths tasqgfshsq
 9121 mtalmsrtpg dvpwlspsv eeassasfsl sspvmtsssp vsstlpdsih ssslpvtsll
 9181 tsglvkttel lgtsssepets sppnlsstsv eilattevtt dteklemtnv vtsgythesp
 40 9241 ssvladsvtt katssmgity ptgdnvltss tpafsdtsri qtskklstlp glmetsisee
 9301 tssatekstv lssvptgatt evsrteaiss srtsipgpaq stmsdtsme titristplt
 9361 rkestdmait pktgpggats qgtftldssss taswpghsa ttqrfpqsvv ttpmsrgped
 9421 vswpsplsv knspssslsv sssvtspspl ystpsgsshs spvpvtslft simmkatdml
 9481 daslepetts apnmnitsde slatskatte teaihvfent aashvettss teelyssspg
 45 9541 fseptkvisp vvtssirdn mvsttmppgss gitrieiesm ssltpglret rtsqdtstt
 9601 etstvlkms sgatpevsrt evmpssrtsi ppgaqstmsl disdevvtrl stspimtesa
 9661 eitittgtgy slatsqvltl lgtstfllsg thstmsqgls hsemnlmsr gpeslswtsp
 9721 rfvettssss sltslpltt lsvsstltd sspssplpvt slilpglvkt tevldtssep
 9781 ktssspnlss tsveipatse imtdtekih sntavakvr tsssvshes svladsetti
 9841 tipsmgitsa vddttvftsn pafsetrrip teptfsltpg fretstseet tsitetsavl
 50 9901 ygvptsatte vsmteimssn rthipdsdqs tmspditev itrllssmm sestqmtitt
 9961 qksspgataq stltlattta plarthstvp prflhsemmt lmsrpenps wksspfvekt
 10021 sssssllslp vttspvsst lpqsspsf svtslltpgm vktdtdstep gtslspnlsg
 10081 tsveilaase vtdtekih pssmavtnvg ttssghelys svsihsepsk atypvgtpss
 55 10141 maetsistm panfettgfe aepfshltsg frktnmsldt ssvtpntps spgsthlq
 10201 sktdftssak tsspdwppas qyteipvdii tpfnaspsit estgitsfpe srftmsvtes
 10261 thhlstdllp saetistgtv mpalsseamts fattgvprai sgsgspfsrt esgpgdatls
 10321 tiaeslpsst pvpfssstft ttdsstipal heitsssatp yrvdtslge ssttegrlvm
 10381 vstldtssqp grtssstpild trmtesvelg tvtsayqvpv lstrlrrtdg imehitkipn
 10441 eaahrgtirp vkpgqtstsp aspklhtgg tkrmetttta lktttalkt tsratltsv
 60 10501 ytptlgtltp lnasrqmast iltemmittp yvfpdvpett sslatslgae tstalprtt
 10561 svlnresett aslvrsrgae rspviqtlav sssepdtas wvhipaetip tvskttnpff

	10621	hseldtvssst	atshgadvss	aipnispse	ldaltplvti	sgtdtsttff	tltkspmete
	10681	trttwlthpa	etsstiprti	pnfshhesda	tpsiatppga	etssaipimt	vspgaedlvt
	10741	sqvtssgtdr	nmtiptltls	pgepktiasl	vthpeaqts	aiptstispa	vsrlvtsmvt
5	10801	slaaktsttn	raltnsppgep	attvslvthp	aqtsptvpwt	tsiffhsksd	ttpsmttshg
	10861	aesssavptp	tvstevpgvv	tplvtssrav	isttipiltl	spggepettps	matshgeeas
	10921	saipptvsvp	gvpgvvtslv	tssravtstt	ipiltfslge	pettpsmats	hgteagsavp
	10981	tvlpvpgmv	tslvassrav	tsttlptltl	spggepettps	matshgaeas	stvpvtvspev
	11041	pgvvtslvts	ssgvnstsip	tlilspgele	ttpsmatshg	aeassavptp	tvspgvsgvv
	11101	tplvtssrav	tsttipiltl	sssepetteps	matshgveas	savltvspv	pgmvstlvt
10	11161	sravtsttip	tlitissdepe	tttstlvthse	akmisaiptl	avspvtvgglv	tslvtssgse
	11221	tsafsnltva	ssqpetidsw	vahpgteass	vvptltvstg	epftnislvt	hpaessstlp
	11281	rttsrfshse	ldtmpstvt	peaesssa	ttispgipgv	ltslvtssgr	disatfptvp
	11341	espheseata	swvthpavts	ttvprttgny	shsepdttps	iatspgaeat	sdfpittvsv
	11401	dvpdmvtsgv	tssgtdtsit	iptltlssge	petttsfity	sethtssaip	tlpvspgask
15	11461	mltslvisg	tdstttfptl	tetpyepett	aiqlihpaet	ntmvpkttk	fshsksdttl
	11521	pvaitspgpe	assavsttti	spdmtdlvt	lvpssgtdts	ttfptlsetp	yepettvtwl
	11581	thpaetsttv	sgtipnfshr	gsdtapsmvt	spgvdrsgv	pttippssip	gvvtssqvtss
	11641	atdtstaipt	ltpspgepet	tassathpgt	qtgftvpipt	vpsssepdtma	swvthppqts
	11701	tpvsrttssf	shsspdattv	matssprteas	savlttispg	apemvtsqit	ssgaatsttv
20	11761	ptlthspgmp	ettallsthp	rtgtsktfpa	stvfppqvset	tasltirpga	etstaltptgt
	11821	tsslftllvt	gtsrvdlst	aspgvsakta	plsthpget	stmipsttll	lglltettgll
	11881	atsssaetst	stltltvspa	vsglssasit	tdkpgtvtsw	ntetspsvts	vgppefsrtv
	11941	tgttmtlips	emptppktsh	gegvspttil	rttmveatnl	attgssptva	kttttfnfntla
	12001	gslftpltt	gmstlasesv	tsrtsynhrs	wisttssynr	rywtpatst	vtstfsgpis
25	12061	tssipsstaa	tvpfmvpftl	nftitnlqye	edmrhpgsrk	fnaterelqg	llkplfrnss
	12121	leylysgcrl	aslrepekdss	amavdaicth	rpdpedlgld	rerlywelsn	ltngiqelgp
	12181	ytldrnslyv	ngfthrssmp	ttstpgtstv	dvgtsgtpss	spspaaagpl	lmpftlnfti
	12241	nlqyeedmr	rtgsrkfntm	esvlqgllkp	lfkntsvgpl	ysgcrltllr	pekdgatgv
	12301	daicthrlp	kspglnreql	ywelsklntd	ieelgpytld	rnslyvngft	hsssvsttst
30	12361	pgtstvdirt	sgtpssllsp	timaagpllv	pftlnftitn	lqygdmghp	gsrkfntter
	12421	vlqgllgpif	kntsvgplys	gcrltstlrse	kdgaatgvda	icihhldpks	pglnrerlyw
	12481	elsqltngik	elgpytldrn	slyvngfthr	tsvpttstpg	tstvdltgts	tpfslpspat
	12541	agpdlvlftl	nftitnlkye	edmrhpgsrk	fnntervlqt	llgpmfknts	vgllysgcrl
	12601	tlrsekdga	atgvdaicth	rlldpkspgld	reqlywelsq	ltngikelgp	ytldrnslyv
35	12661	ngfthwipvp	tsstpgtstv	dlsgtspssl	psptaagpll	vpftlnftit	nlqyeedmh
	12721	pgsrkfntte	rvlqgllgpm	fkntsvglly	sgcrltllrs	ekdgatgvda	aicthrlpdk
	12781	spgvdreql	welsqltngi	kelgpytldr	nslyvngfth	qtsapntstp	gtstvdltgs
	12841	gtpsslpspt	sagpdlvpft	lnftitnlqy	eedmrhpgsr	kfntervlq	gllkplfkst
	12901	svgplysgcr	ltllrsekdg	aatgvdaic	hrlldpkspgv	dreqlywels	qntngikelg
40	12961	pytldrnsly	vngfthqtsa	pntstpgtst	vdltsgtstps	slpsptsagp	llvpftlnft
	13021	itnlqyeedm	hhpgsrkfnt	tervlqgllg	pmfkntsvgl	lysgcrltll	rpekngaag
	13081	mdaicshrl	pkspglnreq	lywelsqlth	gikelgpytl	drnslyvngf	thrsvapts
	13141	tpgtstvdlg	tsgtspsslps	pttavpllv	ftlnftitnl	qygdmrhp	srkfntterv
	13201	lqgllgplfk	nssvgplysg	crlislrsek	dgaatgvda	cthhlnpqp	gldreqlwyq
45	13261	lsqmtngike	lgpytldrns	lyvngfthrs	sglttstpwt	stvdltsgt	pspvpstpta
	13321	gpdlvpftln	ftitnlqyee	dmhrpgsrkf	nttervlqgl	lspifknssv	gplysgcrlt
	13381	slrpekdga	tgmdavclyh	pnpkrpgldr	eqlywelsql	thnitelgpy	sldrdslyvn
	13441	gfthqnsvpt	tstpgtstvy	wattgtppss	pghtepgpl	ipftfnftit	nlhyeenmqh
	13501	pgsrkfntte	rvlqgllkpl	fkntsvgply	sgcrltslrp	ekdgatgmd	avclyhpnk
50	13561	rpglrdreql	welsqlthni	telgpysldr	dslyvngfth	qnsvpttstp	gtstvywatt
	13621	gtpssfpght	epgpdlipft	fnftitnlhy	eenmqhpgsr	kfntervlq	gllkplfknt
	13681	svgplysgcr	ltllrpeke	aatgvdtict	hrvdipgpl	drerlywels	qltnsitelg
	13741	pytldrdsly	vngfnprssv	pttstpgtst	vhlatsgtps	slpghtapvp	llipftlnft
55	13801	itnlhyeenm	qhpgsrkfnt	tervlqgllk	plfkntsvgp	lysgcrltll	rpekheaag
	13861	vditcthrvd	pigpglxxex	lywelsxltx	xixelgpytl	drxslyvngf	thxxsxpsts
	13921	tpgtstvxg	tsgtspsspx	xtsagpdlvp	ftlnftitnl	qyeedmhpg	srkfntterv
	13981	lqgllgpmfk	ntsvgllysg	crltllrpek	ngaagmdai	csrlldpks	gldreqlwy
	14041	lsqlthgike	lgpytldrns	lyvngfthrs	svaptstpgt	stvdltsgt	psslpsptta
	14101	vpdlvpftln	ftitnlqyge	dmhrpgsrkf	nttervlqgl	lgplfkns	gplysgcrlt
60	14161	slrsekdga	tgvdaicthh	lnpqspgldr	eqlywqlsqm	ngikelgpy	tlrnslyvn
	14221	gfthrssglt	tstpwtstvd	lgtsgtspv	pspttagpl	vpftlnftit	nlqyeedmhr

14281 pgsrkfnate rvlqgllspi fknssvgply sgcrltslrp ekdgaatgmd avclyhpnkp
 14341 rpgldreqlly welsqlthni telgpysldr dslyvngfth qssmtttrtp dtstmhlat
 14401 rtpaslsqpt taspllvlfth inctitnlqy eedmrtrgsr kfntmesvlg gllkplfknt
 14461 svqplysgcr ltllrpkkdg aatgvdaict hrlpdkspgl nreqlywels klndieelg
 5 14521 pytlrnsly vngfthqssv ststpgtst vdlrtsgtps slsptimxx xpllxpftxn
 14581 xtitnlxxxx xmxpgsrkf nttervlggl lrplfkntsv sslysgcrlt llrpekdgaa
 14641 trvdaactyr pdpkspglldr eqlywelsql thsitelgpy tldrsvlyvn gfnprssvpt
 14701 tstpgtstvh latsgtpssl pghtxxxppl xpftxxtit nlxxxxmxx pgsrkfntte
 14761 rvlqgllkpl frnssleyly sgcrslaslrp ekdssamavd aicthrpde dgllderly
 10 14821 welsnltngi qelgpytldr nslyvngfth rssglttstp wtstvdlgts gtpspvpspt
 14881 tagpllvfth lnftitnlqy eedmrhpgsr rfnttervlg glltplfknt svqplysgcr
 14941 ltllrpekqe aatgvdtict hrvdpgpgl drerlywels qltnsitelg pytlrdrdsly
 15001 vngfnpwssv pttstpgtst vhlatsgtps slpghtapvp llipftlnft itdlhyeenm
 15061 qhpgsrkfnt tervlqgllk plfkstsvgp lysgcrltll rpekhaatg vdaictlrl
 15121 ptgpgldrer lywelsqltn svtelgpytl drdslyvngf thrssvpts ipgtsavhle
 15181 tsqtpaslpq htapgpplvp ftlnftitnl qyeedmrhpg srkfstterv lqgllkplfk
 15241 ntsvsslysg crltllrpek dgaatrsvd cthrpdpksp gldrerlywk lsqthgite
 15301 lgpytldrns lyvngfthqs smtttrtpdt stmhlatst paslsqptta spllvftin
 15361 ftitnlryee nmhpgsrkf nttervlggl lrpvfkntsv gplysgcrlt tlgpkkdgaa
 20 15421 tkvdaictyr pdpkspglldr eqlywelsql thsitelgpy tqdrdslyvn gftthrsvpt
 15481 tsipgtsavh letsqtpasl pghtapgppl vpftlnftit nlqyeedmrh pgsrkfntte
 15541 rvlqgllkpl fkstsvgply sgcrlltllr ekrgaatgvd ticthrlpdp npgldreql
 15601 welskltrgi ielgpytldr gslyvngfth rtsvpttstp gtstvdlgts gtpfslpspa
 15661 xxxpplxpft xnttitnlxx xmxmxxpgr kfnttervlg tllgpmfkn svqlysgcr
 25 15721 ltllrsekdg aatgvdaict hrlpdkspgv dreqlywels qltngikelg pytlrnsly
 15781 vngfthwipv ptsstpgtst vdlsgtpps lpspttagpl lvpftlnfti tnlkyeedmh
 15841 cpgrskfntt ervlqslgpl mfknstsvgp ysgcrltllr sekdgatgv daicthrlpdp
 15901 kspgvdreql ywelsqltng ikelgpytld rnslyvngft hqtsapntst pgtstvdltg
 15961 sgtpsslpsp txxpplxpft txnttitnlx xxxmxxpgr rkfnttexvl qgllxpfxkn
 30 16021 xsvgxlysgc rltxlrxekx gaatgxdaic xhxxpkxpg lxxexlywel sxtlxxixel
 16081 gpytldrsl yvngfthwip vptsstpgts tvdlsgtpps slpspttagp llvpftlnft
 16141 itnlkyeedm hcpgsrkfnt tervlqslg pmfknstsvgp lysgcrltsl rsekdgatgv
 16201 vdaicthrvd pkspgvdreq lywelsqltn gikelgpytl drnslyvngf thqtsapnts
 16261 tpgtstvxg tsqtpssxpx xtsagpllv ftlnftitnl qyeedmhpg srkfntterv
 35 16321 lqgllgpmfk ntsvlglysg crltllrpek ngattgmdei cthrlpdksp glxxexlywe
 16381 lsxltxxixe lgpytldrsl lyvngfthxx sxpttstpgt stvxxgtstg pssxpxttxx
 16441 xpplxpftxn xtitnlxxxx xmxpgsrkf nttervlggl lkplfrnssl eylysgcrla
 16501 slrpekdsaa mavdaicthr pdpedglldr erlywelsnl tngiqelgpy tldrnslyvn
 16561 gftthrsmpt tstpgtstvd vgtsgtpss pspttagpl ipftlnftit nlqyedmg
 40 16621 pgsrkfntte rvlqgllgpi fkntsvgply sgcrlltslrs ekdgaatgvd aicihhldpk
 16681 spglrnrerly welsqltngi kelgpytldr nslyvngfth rtsvpttstp gtstvdlgts
 16741 gtpfslpspa tagpllvlfth lnftitnlky eedmrhpgsr kfnttervlg tllgpmfkn
 16801 svgllysgcr ltllrsekdg aatgvdaict hrlpdkspgl xxexlywels xltxxixelg
 16861 pytlrdrsl yvngfthxxx pttstpgtst vxxgtsgtps sxpxttxxp llxpftxnt
 45 16921 itnlxxxxxm xpgsrkfnt tervlqgllr pvfkntsvgp lysgcrltll rpkkdgaatk
 16981 vdaictyrpd pkspgldreq lywelsqlth sitelgpytq drdslyvngf thrssvpts
 17041 ipgtsavhle ttgtpssfp htepgpllip ftfnftitnl ryeenmqhpg srkfntterv
 17101 lqglltplfk ntsvlglysg crltllrpek qeaatgvdti cthrvdpigp gldrerlywe
 17161 lsqtnsite lgpytldrds lyvdgfnps svpttstpgt stvhlatsgt psplpghtap
 50 17221 vpllpftln ftitdlhyee nmqhgprkf nttervlggl lkplfkstsv gplysgcrlt
 17281 llrpekhaa tgvdaiactlr ldptgpgldr erlywelsql tnsitelgpy tldrdslyvn
 17341 gfnpwssvpt tstpgtstvh latsgtpssl pghttagpl vpftlnftit nlkyeedmh
 17401 pgsrkfntte rvlqslhgpm fkntsvgply sgcrlltslrs ekdgaatgvd aicthrlpdk
 17461 spglxxexly welsltxxi xelgpytldr xslyvngfth xxxxpttstp gtstvxgts
 55 17521 gtpssxpxt xxxpplxpft xnttitnlxx xmxmxxpgr kfnttexvlq gllxpfxknx
 17581 svgxlysgcr ltxlrxekxg aatgxdaicx hxxpkxpgl xxexlywels xltnsitelg
 17641 pytlrdrdsly vngfthrrsm pttipgtsa vhlatsgtpa slpghtapp llvpftlnft
 17701 itnlqyeedm rhpgsrkfnt tervlqgllk plfkstsvgp lysgcrltll rpekrgaatg
 17761 vdticthrlp plnpglxxex lywelsltx xixelgpytl drxslyvngf thxxsppts
 60 17821 tpgtstvxg tsqtpssxpx txxpplxp ftxnttitnl xxxmxxpgr srkfnttexv
 17881 lqgllxpfxk nxsvgxlysg crltxlrxek xgaatgxdaic cxhxxpkxpg glxxexlywe

17941 lsxltxxixe lgpytldrxx lyvngfhprs svpttstpgt stvhlatsgt psslpghatap
 18001 vpllipftln ftitnlhyee nmqhggsrkf nttervlqgl lgpmpkntsv gllysgcrlt
 18061 llrpekngaa tgmdaicshr ldpkspglxx exlywelsxl txxixelgpy tldrxxslyvn
 18121 gftthxxsxt tstpgtstvx xgtsgtpssx pxxtxxxpll xpftxnxtit nlxxxxmxx
 5 18181 pgsrkfntte xvlqglxpx fknxsvgxly sgcrltxlrx ekxgaatgxd aicxhxxxp
 18241 xpglxxexly welsxltxxi xelgpytldr xslyvngfth qnsvpststpgt gtstvywatt
 18301 gtpssfpgh tpgppllipft fnftitnlhy eenmqhggsr kfnttervlq glltplfknt
 18361 svgplysgcr ltllrpekqe aatgvdtict hrvdipgpgl xxexlywels xltxxixelg
 18421 pytldrxxsly vngfthxxsx pttstpgtst vxxgtsgtps sxxpaxttxp lxpftxnxt
 10 18481 itnlxxxxxm xpgsrkfnt texvlqglx pxfkxsvgx lysgcrltxl rxxekxgaatg
 18541 xdaicxhxxx pkxpglxxex lywelsxltx xixelgpytl drxslyvngf thrssvptts
 18601 spgtstvhla tsqtpsslpq htapvpplip ftlnftitnl hyeenmqhpg srkfntterv
 18661 lqglkplfk stsvgplysg crltllrpek hgaatgvda ctilrdptgp glxxexlywe
 18721 lsxltxxixe lgpytldrxx lyvngfthxx sxpttstpgt stvxxgtsgt pssxpaxttx
 15 18781 xpllxpftxn xtitnlxxxx xmxpgsrkf nttxvlqgl lxpfxkxsv gxlysgcrlt
 18841 xlrxxkxgaa tgdaicxhx xpkxpglxx exlywelsxl txxixelgpy tldrxxslyvn
 18901 gftthrtsvpt tstpgtstvh latstgtpss pghtapvppl ipftlnftit nlqyeedmhr
 18961 pgsrkfntte rvlqglspi fknssvgply sgcrtslrp ekdgaatgmd avclyhpnk
 19021 rpgldreqly celsqlthni telgpytdr dslyvngfth qnsvpststpgt gtstvywatt
 20 19081 gtpssfpgh txxpplxpft xntitnlxx xxxmxxpgsr kfnttexvlq glxpfxkx
 19141 svxlysgcr ltxlrxxkxg aatgxdaicx hxxpxkxpgl xxexlywels xltxxixelg
 19201 pytldrxxsly vngfthwssg lttstptwt vdlgtsgtps pvpspttagp llvpftlnft
 19261 itnlqyeedm hrpgsrkfna tervlqglls pifkntsvgp lysgcrltll rpekqeaatg
 25 19321 vdtictthr vdtictthr vdtictthr lywelsxltx xixelgpytl drxslyvngf thxxsxppts
 19381 tpgtstvx xgtsgtpssx xtxxpllx ftxnxtitnl xxxmxxpgsr srkfnttexv
 19441 lqglxpxfk nxsvgxlysg crltxlrxek xgaatgx dai cxhxxxp kx glxxexlywe
 19501 lsxltxxixe lgpytldrxx lyvngfthrs fglttstptw stvdlgtsgt pspvpsptta
 19561 gpplvpftln ftitnlqyee dmhrpgsrkf nttervlqgl ltpftrntsv sslysgcrlt
 30 19621 llrpekdgaa trvdavcthr pdpkspglxx exlywelsxl txxixelgpy tldrxxslyvn
 19681 gftthxxsxt tstpgtstvx xgtsgtpssx pxxtxxxpll xpftxnxtit nlxxxxmxx
 19741 pgsrkfntte xvlqglxpx fknxsvgxly sgcrltxlrx ekxgaatgxd aicxhxxxp
 19801 xpglxxexly welsxltxxi xelgpytldr xslyvngfth wipvptsstpgt gtstvdllgsg
 19861 tpsslpstt agpplvpftl nftitnlqyg edmhpgsrkf fnttervlqg llgpifknt
 35 19921 vgpysgcrl tslrsekdg aatgvdaic hldpkspglx xexlywelsx ltxixelg
 19981 ytldrxxsly vngfthxxsx tttstpgtstv xgtsgtpss xpxttxxpl lxpftxnxti
 20041 tnlxxxxmxx xpgsrkfntt exvlqglxpx fknxsvgxly ysgcrltxlr xekxgaatgx
 20101 daicxhxxxp kxpglxxex lywelsxltx ixelgpytld rxslyvngft hqtfaantst
 20161 pgtstvdltg sgtppslpsp tsagpllvf tlnftitnlq yeedmhpgs rkfnttervl
 40 20221 qglgpmfkn tsvgllysgc rltllrpekn gaatr vdvac thrpdkspg lxxexlywel
 20281 sxltxxixel gpytldrxxsly vngfthxxsx xpttstpgts tvxxgtsgt sxxpaxtapv
 20341 pllipftlnf titnlhyeen mqhggsrkfn ttervlqgl kplfkstsvg plysgcrltl
 20401 lrpekhaat gvdaictlrl dptgpgldre rlywelsqlt nsvtelgpyt ldrdslyvng
 20461 ftqrssvptt sipgtsavhl etsgt paslp ghtapgpplv pftlnftitn lqyevdmrhp
 45 20521 gsrkfntter vlqglkplf kstsvgplys gcrltllrpe krgaatgvdt icthrldpln
 20581 pgldreqlyw elskltrgii elgpyldrg slyvngfth nfvpitstpg tstvhltgse
 20641 tpsslprpiv pgpllvpftl nftitnlqye eamrhpgsrkf fnttervlqg llrplfknts
 20701 igplysscr lllrpekdk aatr vdaicth hpdqspgln reqlywelsq lthgitelgp
 20761 ytdrdslyv dgfthwspip tttstpgtsiv nlgtsgipps lpettxxxpl lxpftxnxti
 50 20821 tnlxxxxmxx xpgsrkfntt ervlqglkpx lfkstsvgpl ysgcrltllr pekdgvatrv
 20881 daicthrpdp kipgldrqgl ywelsqlths itelgpytld rdslyvngft qrssvpttst
 20941 pgtftvqpet setpsslpqp tatgpvllpf tlnftitnlq yeedmhpgs rkfnttervl
 21001 qglmlplfk tsvsslysgc rltllrpekd gaatr vdvac thrpdkspg ldrerlywkl
 21061 sqldthgitel gpytldrxxsly vngfthqss mtttrtpdts tmhlatsrtp aslsgpttas
 55 21121 pllvftinf titnlryeen mhhpgsrkf ttervlqgl rpvfkntsvg plysgcrltl
 21181 lrpkkdgaat kvdaictyrp dpkspgldre glywelsqlt hsitegpyt ldrdslyvng
 21241 ftqrssvptt sipgtptvdl gtsgtpvskp gpasaplly lftlnftitn lryeenmqhp
 21301 gsrkfntter vlqglrslf kstsvgplys gcrltllrpe kdgtatgvda ictthpdpks
 21361 prldreqlyw elsqldthni elghyaldnd slfvngfth ssvsttstpg rtpvylgask
 60 21421 tpassifgsa ashllilftl nftitnlrye enmwpgsrkf nttervlqgl lrpplfkntsv
 21481 gpysgsrlyt llrpekdgaa tgvdacithr pdptgpgldr eqlylelsq lthsitegpy
 21541 tldrdslyvn gftthrsvpt tstgvvseep ftlnftinnl rymadmgpgg slkfnitdnv

21601 mkhlslsplfq rsslrgarytg crvialrsvk ngaetrvdll ctylqplsgp glpikqvfh
 21661 lsqgthgitr lgpysldkds lyngynepg ldeppptpkp attflpplse attamgyhlk
 21721 tltlnftisn lqyspdmkgk satfnstegv lqhlrlplfq kssmgpfylg cqlislrpek
 21781 dgaatgvdt ctlyhdpvvp gldiqqllywe lsqllthgvtq lgfyvldrds lfingyapqn
 5 21841 lsirgeyqin fhivnwnlsn pdptsseyit llrdiqdkvt tlykgsqldh tfrfclvtnl
 21901 tmdsvlvtvk alfssnldps lveqvfldkt lnasfhwlg tyqlvdihvt emessvyqpt
 21961 sssstqhfyf nftitnlpys qdkaqpgttn yqrnkknied alnqlfrnss iksyfsdcqv
 22021 stfrsvpnrh htgvdslnnf splarrvdrv aiyeeflrmr rngtqlqnft ldrssvldvg
 22081 yspnrneplt gnsdlpfwav iliglagllg litclicgvl vttrrrkkek eynvqqcpg
 10 22141 yyqshldled lq

SEQ ID NO. 2

15 CA125 nucleic acid Genbank No. AF414442

CDS 205..66663

1 aagcgttgca caattccccc aacctccata catacggcag ctcttctaga cacaggtttt
 61 cccaggtcaa atgcggggac cccagccata tctcccaccc tgagaaattt tggagtttca
 121 gggagctcag aagctctgca gaggccaccc tctctgaggg gattcttctt agacctccat
 20 181 ccagaggcaa atgttgacct gtccatgctg aaacctcag gccttcctgg gtcatcttct
 241 cccaccgct ccttgatgac agggagcagg agcactaaag ccacaccaga aatggattca
 301 ggactgacag gagccacctt gtcacctaaag acatctacag gtgcaatcgt ggtgacagaa
 361 catactctgc cctttacttc cccagataag accttgcca gtctacatc ttcggttggtg
 421 ggaagaacca cccagtcttt ggggtgatg tctctgctc tccctgagtc aacctctaga
 25 481 ggaatgacac actccgagca aagaaccagc ccacgtgta gtccccagggt caatggaaact
 541 ccctctagga actaccctgc tacaagcatg gtttcaggat tgagtcccc aaggaccagg
 601 accagttcca cagaaggaaa tttaccaaa gaagcatcta catacacact cactgtagag
 661 accacaagtg gcccagtcac tgagaagtac acagtcccca ctgagacctc aacaactgaa
 721 ggtgacagca cagagacccc ctgggacaca agatatattc ctgtaaaaat cacatctcca
 30 781 atgaaaacat ttgcagattc aactgcatcc aaggaaaatg cccagtgctc tatgactcca
 841 gctgagacca cagttactga ctcacatact ccaggaagga caaacccatc atttgggaca
 901 ctttattctt ccttccttga cctatcacct aaagggaacc caaattccag aggtgaaaca
 961 agcctggaac tgattctatc aaccttgga tatcccttct cctctcctga acctggctct
 1021 gcaggacaca gcagaataag taccagtgcg cctttgtcat catctgcttc agttctcgat
 35 1081 aataaaatat cagagaccag catattctca ggccagagtc tcacctcccc tctgtctcct
 1141 ggggtgcccg agggcagagc cagcacaatg cccaactcag ctatcccttt ttccatgact
 1201 ctaagcaatg cagaaaacaag tgccgaaagg gtcagaagca caatttctct tctggggact
 1261 ccatcaatat ccacaaagca gacagcagag actatcctta ccttcctatgc cttcgctgag
 1321 accatggata taccagcac ccacatagcc aagactttgg cttcagaatg gttgggaagt
 40 1381 ccaggtaccc ttggtggcac cagcacttca gcgctgacaa ccacatctcc atctaccact
 1441 ttagtctcag aggagaccaa caccatcac tccacgagtg gaaaggaaac agaaggaaact
 1501 ttgaatacat ctatgactcc acttgagacc tctgctcctg gagaagagtc cgaatgact
 1561 gccaccttgg tccccactct aggttttaca actcttgaca gcaagatcag aagtccatct
 1621 caggtctctt catcccaccc aacaagagag ctccagaacca caggcagcac ctctgggagg
 45 1681 cagagttcca gcacagctgc ccacgggagc tctgacatcc tgagggcaac cacttccagc
 1741 acctcaaaag catcatcatg gaccagtga agcacagctc agcaatttag tgaacccag
 1801 cacacacagt ggggtggagac aagtccatgc atgaaaacag agagaccccc agcatcaacc
 1861 agtgtggcag cccctatcac cacttctgtt ccctcagtg tctctggctt caccaccctg
 1921 aagaccagct ccacaaaagg gatttggctt gaagaaacat ctgcagacac actcatcgga
 50 1981 gaatccacag ctggcccaac caccatcag tttgctgttc ccactgggat ttcaatgaca
 2041 ggaggcagca gcaccagggg aagccagggc acaaccacc tactcaccag agccacagca
 2101 tcatctgaga catccgcaga tttgactctg gccacgaacg gtgtccagc tccgtgtct
 2161 ccagcagtga gcaagacggc tgctggctca agtcctccag gagggacaaa gccatcatat
 2221 acaatggttt cttctgtcat cctgagaca tcatctctac agtcctcagc tttcagggaa
 55 2281 ggaaccagcc tgggactgac tccattaaac actagacatc ccttctcttc cctgaacca
 2341 gactctgcag gacacaccaa gataagcacc agcattcctc tgttgcac tgcctcagtt
 2401 cttgaggata aagtgtcagc gaccagcaca ttctcacacc acaaagccac ctcatctatt
 2461 accacagggg ctctgaaat ctcaacaaag acaaagccca gctcagccgt tcttctctcc
 2521 atgaccctaa gcaatgcagc acaagtcct gaaagagtca gaaatgcaac ttccctctg

	2581	actcatccat	ctccatcagg	ggaagagaca	gcagggagtg	tcctcactct	cagcacctct
	2641	gctgagacta	cagactcacc	taacatccac	ccaactggga	cactgacttc	agaatcgtca
	2701	gagagtcccta	gcactctcag	cctcccaggt	gtctctggag	tcaaaaccac	atcttcttca
	2761	tctactcctt	ccactcatct	atttactagt	ggagaagaaa	cagaggaaac	ttcgaatcca
5	2821	tctgtgtctc	aacctgagac	ttctgtttcc	agagtaagga	ccaccttggc	cagcacctct
	2881	gtccctaccc	cagtattccc	cacctatggac	acctggccta	cacgttcagc	tcagttctct
	2941	tcatcccacc	tagtgagtga	gctcagagct	acgagcagta	cctcagttac	aaactcaact
	3001	ggttcagctc	ttcctaaaat	atctcacctc	actgggacgg	caacaatgtc	acagaccaat
10	3061	agagacacgt	ttaatgactc	tgctgcaccc	caaagcacaa	cttggccaga	gactagttcc
	3121	agattcaaga	cagggttacc	ttcagcaaca	accactgttt	caacctctgc	cacttctctc
	3181	tctgtacttg	taatgggtctc	taaattcact	tctccagcaa	ctagttccat	ggaagcaact
	3241	tctatcaggg	aaccatcaac	aacctcctc	acaacagaga	ccacgaatgg	cccaggctct
	3301	atggctgtgg	cttctaccaa	catcccatt	ggaaaagggt	acattactga	aggaagattg
	3361	gacacaagcc	atctgcccac	tggaaaccaca	gcttctctg	agacatctat	ggattttacc
15	3421	atggccaaag	aaagtgtctc	aatgtcagta	tctccatctc	agtcacatga	tgctgtctgg
	3481	tcaagcactc	caggaaggac	aagccaattc	gttgacacat	tttctgatga	tgtctatcat
	3541	ttaacatcca	gagaaattac	aatacctaga	gatggaaaca	gctcagctct	gactccacaa
	3601	atgactgcaa	ctcacctccc	atctcctgat	cctggctctg	ctagaagcac	ctggcttggc
	3661	atcttgtcct	catctccttc	ttctcctact	cccaaagtca	caatgagctc	cacattttca
20	3721	actcagagag	tcaccacaag	catgataatg	gacacagttg	aaactagtcg	gtggaacatg
	3781	cccaacttac	cttccacgac	ttccctgaca	ccaagttaata	ttccaacaag	tggtgccata
	3841	ggaaaaagca	ccctgggttcc	cttggacact	ccatctccag	ccacatcatt	ggaggcatca
	3901	gaagggggac	ttccaaccct	cagcacctac	cctgaatcaa	caaacacacc	cagcatccac
	3961	ctcggagcac	acgctagtct	agaaaagtcca	agcaccatca	aacttaccat	ggcttcagta
25	4021	gtaaaacctg	gctcttacac	acctctcacc	ttcccctcaa	tagagaccca	cattcatgta
	4081	tcaacagcca	gaatggctta	ctcttctggg	tcttcacctg	agatgacagc	tcctggagag
	4141	actaacactg	gtagtacctg	ggaccccacc	acctacatca	ccactacgga	tcctaaggat
	4201	acaagtccag	ctcagggtctc	tacaccctac	tcagtggaga	cactcagaac	acagaaaaac
	4261	catccaaaga	cagagtccgc	caccccagct	gcttactctg	gaagtccata	aatctcaagt
30	4321	tcaccaaatc	tcaccagtcc	ggccacaaaa	gcatggacca	tcacagacac	aactgaacac
	4381	tcactcaat	tacattacac	aaaattggga	gaaaaatcat	ctggatttga	gacacagtca
	4441	gctccaggac	ctgtctctgt	agtaatccct	acctccccta	ccattggaag	cagcacattg
	4501	gaactaactt	ctgatgtccc	aggggaaccc	ctggtccttg	ctccagtga	gcagaccaca
35	4561	atcactctcc	ccatggcaac	atggctgagt	accagtttga	cagaggaaat	ggcttcaaca
	4621	gaccttgata	tttcaagtcc	aagttcaccc	atgagtacat	ttgtattttt	tcacctatg
	4681	tcacacacct	ctcatgaact	ttcaaagtca	gaggcagata	ccagtgcctt	tagaaataca
	4741	gattcaacaa	cggtggatca	gcaccttaga	atcaggagtt	tgggcagaac	tggggactta
	4801	acaactgttc	ctatcacccc	actgacaacc	acgtggacca	gtgtgattga	acactcaaca
40	4861	caagcacagg	acaccttttc	tgcaacgatg	agtcctactc	acgtgacaca	gtactcaaaa
	4921	gatcaaacat	ctataccagc	ctcagcatcc	ccttcccata	ttactgaagt	ctacctgag
	4981	ctcgggacac	aaggggagaag	ctcctctgag	gcaaccactt	tttggaaaac	atctacagac
	5041	acactgtcca	gagagattga	gactggccca	acaaacattc	aatccactcc	acccatggac
	5101	aacacaacaa	cagggagcag	tagtagtgga	gtcacccctg	gcatagccca	ccttccata
45	5161	ggaacatcct	ccccagctga	gacatccaca	aacatggcac	tggaaagaag	aagttctaca
	5221	gccactgtct	ctatggctgg	gacaatggga	ctccttggtt	ctagtgtctc	aggaagaagc
	5281	atcagccagt	cattaggaag	agtttctctc	gtcctttctg	agtcactac	tgaaggagtc
	5341	acagattcta	gtaaggggaag	cagcccaagg	ctgaacacac	agggaaatac	agctctctcc
	5401	tcctctcttg	aaccagctca	tgctgaaggc	agccagatga	gcacaagcat	ccctcaacc
	5461	tcactctcta	caactcctga	tgtggaattc	atagggggca	gcacattttg	gaccaaggag
50	5521	gtcaccacag	ttatgacctc	agacatctcc	aagtcttcag	caaggacaga	gtccagctca
	5581	gctaccctta	tgtccacagc	tttgggaagc	actgaaaata	caggaaaaga	aaaactcaga
	5641	actgcctcta	tggatcttcc	atctccaact	ccatcaatgg	aggtgacacc	atggattttc
	5701	ctcactctca	gtaatgcccc	caataccaca	gattcacttg	acctcagcca	tggggtgcac
55	5761	accagctctg	cagggacttt	ggccactgac	aggtcattga	atactggtgt	cactagagcc
	5821	tcagatttgg	aaaacggctc	tgatacctct	tctaagtccc	tgtctatggg	aaacagcact
	5881	cacacttcca	tgactgacac	agagaagagt	gaagtgtctt	cttcaatcca	tcccggacct
	5941	gagacctcag	ctcctggagc	agagaccact	ttgacttcca	ctcctggaaa	cagggccata
	6001	agcttaacat	tgcccttttc	atccattoca	gtggaagaag	tcattttctac	aggcataacc
60	6061	tcaggaccag	acatcaactc	agcaccctatg	acacattctc	ccatcacccc	accaacaatt
	6121	gtatggacca	gtacaggcac	aattgaacag	tccactcaac	cactacatgc	agttttctca
	6181	gaaaaagttt	ctgtgcagac	acagtcaact	ccatatgtca	actctgtggc	agtgctgtct

	6241	tccctaccc	atgagaattc	agtctcttct	ggaagcagca	catcctctcc	atattcctca
	6301	gcctcacttg	aatccttgga	ttccacaatc	agtaggagga	atgcaatcac	ttcctggcta
	6361	tggggacctca	ctacatctct	ccccactaca	acttggccaa	gtactagttt	atctgaggca
	6421	ctgtcctcag	gccattctgg	ggtttcaaac	ccaagttcaa	ctacgactga	atttccactc
5	6481	ttttcagctg	catccacatc	tgctgctaag	caaagaaatc	cagaaacaga	gacccatggg
	6541	ccccagaata	cagccgcgag	tactttgaac	actgatgcac	cctcggctac	aggtctttct
	6601	gagactcctg	tgggggcaag	tatcagctct	gaagtccctc	ttccaatggc	cataacttct
	6661	agatcagatg	tttctggcct	tacatctgag	agtactgcta	acccgagttt	aggcacagcc
	6721	tcttcagcag	ggaccaaatt	aactaggaca	atatccctgc	ccacttcaga	gtccttgggt
10	6781	tccttttagaa	tgaacaagga	tccatggaca	gtgtcaatcc	ctttgggggc	ccatccaact
	6841	actaatagag	aaacaagcat	cccagtaaac	agcgcaggtc	cacctggctt	gtccacagta
	6901	gcacatcagatg	taattgacac	accttcagat	ggggctgaga	gtattccac	tgtctccttt
	6961	tcccctccc	ctgatactga	agtgaacat	atctcacatt	tcccagaaaa	gacaactcat
	7021	tcattttagaa	ccatttcac	tctcactcat	gagttgactt	caagagtgac	acctattcct
15	7081	ggggattgga	tgagttcagc	tatgtctaca	aagcccacag	gagccagtc	ctccattaca
	7141	ctgggagaga	gaaggacaat	cacotctgct	gctccaacca	cttccccc	agttctcact
	7201	gctagtttca	cagagaccag	cacagtttca	ctggataatg	aaactacagt	aaaaacctca
	7261	gatatacctg	acgcacggaa	aacaaatgag	ctcccctcag	atagcagttc	ttctctctgat
	7321	ctgatcaaca	cctccatagc	ttcttcaact	atggatgtca	ctaaaacagc	ctccatcagt
20	7381	cccactagca	tctcaggaat	gacagcaagt	tcctccccc	ctctcttctc	ttcagataga
	7441	ccccaggttc	ccacatctac	aacagagaca	aatacagcca	cctctccatc	tgtttccagt
	7501	aacacctatt	ctcttgatgg	gggctccaat	gtgggtggca	ctccatccac	tttaccaccc
	7561	tttacaatca	cccacctgt	cgagacaagc	tcggccctat	tagcctggtc	tagaccagta
	7621	agaactttca	gcaccatggg	cagcactgac	actgcctcag	gagaaaatcc	tacctctagc
25	7681	aattctgtgg	tgacttctgt	tccagcacca	ggtacatggg	ccagtgtagg	cagtactact
	7741	gacttacctg	ccatgggctt	tctcaagaca	agtccctgcag	gagaggcaca	ctcacttcta
	7801	gcacaaacta	ttgaaccagc	cactgccttc	actcccccac	tctcagcagc	agtggtcact
	7861	ggatccagtg	ctacatcaga	agccagtcct	ctcactacga	gtgaaagcaa	agccattcat
	7921	tcttcaccac	agaccccaac	tacacccacc	tctggagcaa	actgggaaac	actgactact
30	7981	cctgagagcc	ttttggtagt	cactgagact	tcagacacaa	cacttacctc	aaagattttg
	8041	gtcacagata	ccatcttggt	ttcaactgtg	tccacgccac	cttctaaatt	tccaagtaag
	8101	gggactctgt	ctggagcttc	cttccctact	ttactcccgg	acactccagc	catccctctc
	8161	actgccactg	agccaacaag	ttcattagct	acatcctttg	attccacccc	actgggtgact
	8221	atagcttcgg	atagtcttgg	cacagtcoca	gagactaccc	tgaccatgtc	agagacctca
35	8281	aatgggtgatg	cactgggtct	taagacagta	agtaacccag	ataggagcat	ccctggaatc
	8341	actatccaag	gagtaacaga	aagtccactc	catccttctt	ccacttcccc	ctctaagatt
	8401	gttgctccac	ggaatacaac	ctatgaaggt	tcgatcacag	tggcactttc	tactttgcct
	8461	gcgggaacta	ctggttccct	tgtattcagt	cagagttctg	aaaactcaga	gacacggct
	8521	ttggtagact	catcagctgg	gcttgagagg	gcactgtgtga	tgccactaac	cacaggagagc
40	8581	cagggtatgg	ctagctctgg	aggaatcaga	agtgggtcca	ctcactcaac	tggaacccaa
	8641	acattttctt	ctctccctct	gacctgaac	ccagggtgag	ttacagccat	gtctgaaatc
	8701	accacgaaca	gactgacagc	tactcaatca	acagcaccca	aagggatacc	tgtgaagccc
	8761	accagtgtg	agtcaggcct	cctaacacct	gtctctgcct	cctcaagccc	atcaaaggcc
	8821	tttgccctcac	tgactacagc	tccccatca	acttggggga	tcccacagtc	taccttgaca
45	8881	tttgagtttt	ctgaggtccc	aagtttggat	actaagtccg	cttctttacc	aactcctgga
	8941	cagtccctga	acaccattcc	agactcagat	gcaagcacag	catcttccct	actgtccaag
	9001	tctccagaaa	aaaacccaag	ggcaaggatg	atgacttcca	caaaggccat	aagtgcaagc
	9061	tcatttcaat	caacagggtt	tactgaaacc	cctgagggat	ctgcctcccc	ttctatggca
50	9121	gggcatgaac	ccagagtcct	cacttcagga	acaggggacc	ctagatatgc	ctcagagagc
	9181	atgtcttatc	cagacccaag	caaggcatca	tcagctatga	catcgacctc	tcttgcatca
	9241	aaactcacia	ctctcttcag	cacaggtcaa	gcagcaaggt	ctggttctag	ttcctctccc
	9301	ataagcctat	ccactgagaa	agaaacaagc	ttcctttccc	ccactgcatt	cacctccaga
	9361	aagacttcaac	tatttcttgg	gccttccatg	gcaaggcagc	ccaacatatt	gggtgatctt
55	9421	cagacttcag	ctctgacact	ttctccaaca	tccactctaa	atatgtccca	ggagagcct
	9481	cctgagttaa	cctcaagcca	gaccattgca	gaagaagagg	gaacaacagc	tgaaacacag
	9541	acgttaacct	tcacaccatc	tgagacccca	acatccttgt	tacctgtctc	ttctccaca
	9601	gaacccacag	ccagaagaaa	gagttctcca	gaaacatggg	caagctctat	ttcagttcct
	9661	gccaagacct	ccttgggtga	aacaactgat	ggaacgctag	tgaccacct	aaagatgtca
	9721	agccaggcag	cacaaggaaa	ttccacgtgg	cctgccccag	cagaggagac	ggggaccagt
60	9781	ccagcaggca	catccccagg	aagcccagaa	gtgtctacca	ctctcaaaat	catgagctcc
	9841	aaggaaacca	gcacagccc	agagatcagg	tccactgtgc	gaaattctcc	ttggaagact

	9901	ccagaaacaa	ctgttcccat	ggagaccaca	gtggaaccag	tcacccttca	gtccacagcc
	9961	ctaggaagtg	gcagcaccag	catctctcac	ctgcccacag	gaaccacatc	accaaccaag
	10021	tcaccaacag	aaaatatgtt	ggctacagaa	agggtctccc	tctccccatc	cccacctgag
5	10081	gcttggacca	acctttattc	tggaaactcca	ggaggggacca	ggcagtcact	ggccacaatg
	10141	tctctgtct	ccctagagtc	accaactgct	agaagcatca	caggggactgg	tcagcaaagc
	10201	agtccagaac	tggtttcaaa	gacaactgga	atggaattct	ctatgtggca	tggctctact
	10261	ggagggacca	cagggggacac	acatgtctct	ctgagcacat	cttccaatat	ccttgaagac
	10321	cctgtaacca	gccc aaactc	tgtgagctca	ttgacagata	aatccaaaca	taaaaccgag
	10381	acatgggtaa	gcaccacagc	cattccctcc	actgtcctga	ataataagat	aatggcgact
10	10441	gaacaacaga	caagtogatc	tgtggatgag	gcttattcat	caactagttc	ttggctcagat
	10501	cagacatctg	ggagtgcacat	caoccttggg	gcatctcctg	atgtcacaaa	cacattatac
	10561	atcacctcca	cagcacaac	caactcacta	gtgtctctgc	cctctggaga	ccaaggcatt
	10621	acaagcctca	ccaatccctc	aggaggaaaa	acaagctctg	cgtcatctgt	cacatctcct
	10681	tcaatagggc	ttgagactct	gagggccaat	gtaagtgcag	tgaaaagtga	cattgcccc
15	10741	actgctgggc	atctatctca	gacttcatct	cctgcggaag	tgagcatcct	ggacgtaacc
	10801	acagctccta	ctccaggtat	ctccaccacc	atcaccacca	tgggaaccaa	ctcaatctca
	10861	actaccacac	ccaacccaga	agtgggtatg	agtaccatgg	acagcaccoc	ggccacagag
	10921	aggcgacaaa	cttctacaga	acacccttcc	acctggctct	ccacagctgc	atcagattcc
	10981	tggactgtca	cagacatgac	ttcaaacttg	aaagtgtgaa	gatctcctgg	aacaatttcc
20	11041	acaatgcata	caacttcatt	cttagcctca	agcactgaat	tagactccat	gtctactccc
	11101	catggccgta	taactgtcat	tggaaaccagc	ctggtcactc	catcctctga	tgcttcagct
	11161	gtaaagacag	agaccagtac	aagtgaagaa	acattgagtc	cttcagacac	aactgcatct
	11221	actcccatct	caactttttc	tctgtctcag	aggatgagca	tctcagttcc	tgacatttta
	11281	agtacaagtt	ggactcccag	tagtacagaa	gcagaagatg	tgcctgtttc	ataagtttct
25	11341	acagatcatg	ctagtacaaa	gactgaccca	aatacgcccc	tgtccacttt	tctgtttgat
	11401	tctctgtcca	ctcttgactg	ggacactggg	agatctctgt	catcagccac	agccactacc
	11461	tcagctcctc	agggggccac	aactccccag	gaactcactt	tggaaaccat	gatcagccca
	11521	gtaccctcac	agttgcccct	ctctataggg	cacattacaa	gtgcagtcac	accagctgca
	11581	atggcaagga	gctctggagt	tactttttca	agaccagatc	ccacaagcaa	aaaggcgag
30	11641	cagacttcca	ctcagcttcc	caccaccact	tctgcacatc	cagggcgagg	gcccagatca
	11701	gcagcaacaa	ctctggatgt	gatccacac	acagcaaaaa	ctccagatgc	aacttttcag
	11761	agacaagggc	agacagctct	tacaacagag	gcaagagcta	catctgactc	ctggaatgag
	11821	aaagaaaaat	caaccccaag	tgccacttgg	atcactgaga	tgatgaattc	tgtctcagaa
	11881	gataccatca	aggagggttac	cagctcctcc	agtgtattaa	aggaccctga	atagctggga
35	11941	cataaaacttg	gaatctggga	cgacttcatc	cccaagtttg	gaaaagcagc	ccatatgaga
	12001	gagttgcccc	ttctgagctc	accacaggac	aaagaggcaa	ttcacccttc	tacaacacac
	12061	gtagagacca	caggctgggt	cacaagttcc	gaacatgctt	ctcattccac	tatcccagcc
	12121	catcagcgt	catccaaact	cacatctcca	gtggttacaa	cctccaccag	ggaacaagca
	12181	atagtttcta	tgtcaacaac	cacatggcca	gagttctaca	gggctagaa	agagccta
40	12241	tccttcttga	ctattgaaat	gagggacgtc	agcccttaca	tggacaccag	ctcaaccaca
	12301	caaacaagta	ttatctcttc	cccaggttcc	actgcgatca	ccaaggggcc	tagaacagaa
	12361	attacctcct	ctaagagaat	atccagctca	ttccttgccc	agtctatgag	gtcgtcagac
	12421	agccctcag	aagccatcac	caggctgtct	aactttcctg	ccatgacaga	atctggagga
45	12481	atgatccttg	ctatgcaaac	aagtccacct	ggcgctacat	cactaagtgc	acacttttg
	12541	gatacatcag	ccacagcctc	ctggacaggg	actccactgg	ctacgactca	gagatttaca
	12601	tactcagaga	agaccactct	ctttagcaaa	ggctcctgag	atacatcaca	gccaagccct
	12661	ccctctgtgg	aagaaaccag	ctcttctctc	tccctgggtac	ctatccatgc	tacaacctcg
	12721	ccttccaata	ttttgttgac	atcacaaggg	cacagtccct	cctctactcc	acctgtgacc
50	12781	tcagttttct	tgtctgagac	ctctggcctg	gggaagacca	cagacatgtc	gaggataagc
	12841	ttggaacctg	gcacaagttt	acctcccaat	ttgagcagta	cagcaggtga	ggcgttatcc
	12901	acttatgaag	cctccagaga	tacaaaggca	attcatcatt	ctgcagacac	agcagtgacg
	12961	aatatggagg	caaccagtcc	tgaatattct	cctatcccag	gccatacaaa	gccatccaaa
	13021	gccacatctc	cattgggttac	ctcccacatc	atgggggaca	tcacttcttc	cacatcagta
	13081	tttggctcct	ccgagaccac	agagattgag	acagtgtcct	ctgtgaacca	gggacttcag
55	13141	gagagaagca	catcccagggt	ggccagctct	gctacagaga	caagcactgt	cattacccat
	13201	gtgtctagtg	gtgatgctac	tactcatgtc	accaagacac	aagccacttt	ctctagcgga
	13261	acatccatct	caagccctca	tcagtttata	acttctacca	acacattttac	agatgtgagc
	13321	accaaccctc	ccacctctct	gataatgaca	gaatcttcag	gagtgaccat	caccacccaa
	13381	acaggtccta	ctggagctgc	aacacagggt	ccatatctct	tggacacatc	aacctgcct
60	13441	tacttgacag	agactccatt	agctgtgact	ccagatttta	tgcaatcaga	gaagaccact
	13501	ctcataagca	aagggtccaa	ggatgtgacc	tggacaagcc	ctccctctgt	ggcagaaacc

	13561	agctatccct	cttccttgac	acctttcttg	gtcacaaacca	tacctcctgc	cacttccacg
	13621	ttacaagggc	aacatacatc	ctctcctggt	tctgcgactt	cagttctttac	ctctggactg
	13681	gtgaagacca	cagatatggt	gaacacaagc	atggaacctg	tgaccaattc	acctcaaaat
	13741	ttgaacaatc	catcaaatga	gatactggcc	acttttgccag	ccaccacaga	tatagagact
5	13801	attcatcctt	ccataaacia	agcagtgaac	aatatgggga	ctgccagttc	agcacatgta
	13861	ctgcattcca	ctctcccagt	cagctcagaa	ccatctacag	ccacatctcc	aatgggtcct
	13921	gcctccagca	tgggggagcg	tcttgcttct	atatcaatac	ctggttctga	gaccacagac
	13981	attgagggag	agccaacatc	ctccctgact	gctggacgaa	aagagaacag	cacctccag
	14041	gagatgaact	caactacaga	gtcaaacatc	atcctctcca	atgtgtctgt	gggggctatt
10	14101	actgaagcca	caaaaatgga	agtccctctc	tttgatgcaa	cattcatacc	aactcctgct
	14161	cagtcaacaa	agttcccaga	tattttctca	gtagccagca	gtagactttc	aaactctcct
	14221	cccagacaaa	tatctaccca	catgaccacc	acccagacag	ggtcttctgg	agctacatca
	14281	aagattccac	ttgccttaga	cacatcaacc	ttggaaacct	cagcagggag	tccatcagtg
	14341	gtgactgagg	gggttgccca	ctcaaaaata	accactgcaa	tgaaccaatga	tgtcaaggac
15	14401	gtgtcacaga	caaaccctcc	ctttcaggat	gaagccagct	ctccctcttc	tcaagcacct
	14461	gtccttgtca	caaccttacc	ttcttctggt	gctttcacac	cgcaatggca	cagtacctcc
	14521	tctcctgttt	ctatgtcctc	agttcttact	tcttctactg	ttaaagaccgc	aggcaagggtg
	14581	gatacaagct	tagaaacagt	gaccagttca	cctcaaagta	tgagcaacac	tttggatgac
	14641	atatcggtca	cttcagcagc	caccacagat	atagagacaa	cgcactcttc	cataaacaca
20	14701	gtagttacca	atgtggggac	caccgggttc	gcatttgaat	cacattctac	tgtctcagct
	14761	taccagagc	catctaaagt	cacatctcca	aatgttacca	cctccaccat	ggaagacacc
	14821	acaatttccc	gatcaatacc	taaatcctct	aagactacaa	gaactgagac	tgagacaact
	14881	tcctccctga	ctcctaaact	gagggagacc	agcatctccc	aggagatcac	ctcgtccaca
	14941	gagacaagca	ctgttcttta	caaagagctc	actgggtgcca	ctaccgaggt	atccaggaca
25	15001	gatgtcactt	cctctagcag	tacatccttc	cctggccctg	atcagtccac	agtgtcacta
	15061	gacatctcca	cagaaaccaa	caccaggctg	tctacctccc	caataatgac	agaatctgca
	15121	gaaataacca	tcaccaccca	aacagggtcct	catggggcta	catcacagga	tacttttacc
	15181	atggacccat	caaatacaac	ccccaggca	gggatccact	cagctatgac	tcattggattt
	15241	tcacaattgg	atgtgaccac	tcttatgagc	agaattccac	aggatgtatc	atggacaagt
30	15301	cctccctctg	tgataaaaac	cagctccccc	tcttctcttc	tgtcctcacc	tgcaatgacc
	15361	acaccttccc	tgatttcttc	taccttacca	gaggataagc	tctcctctcc	tatgacttca
	15421	cttctcacct	ctggcctagt	gaagattaca	gacatattac	gtacacgctt	ggaacctgtg
	15481	accagctcac	ttccaaattt	cagcagcacc	tcagataaga	tactggccac	ttctaagac
	15541	agtaaagaca	caaaggaaat	tttctcttct	ataaacacag	aagagaccaa	tgtgaaagcc
35	15601	aacaactctg	gacatgaatc	ccattcccct	gcactggctg	actcagagac	acccaaagcc
	15661	acaactcaaa	tggttatcac	caccaactgtg	ggagatccag	ctccttccac	atcaatgcca
	15721	gtgcatgggt	cctctgagac	tacaaacatt	aagagagagc	caacatattt	cttgactcct
	15781	agactgagag	agaccagtag	ctctcaggag	tccagcttct	ccacggacac	aagtcttcta
	15841	ctttccaaag	ttcccactgg	tactattact	gagggtctcca	gtacaggggt	caactcttct
40	15901	agcaaaattt	ccaccccaga	ccatgataag	tcacagtgcc	cacctgacac	cttcacagga
	15961	gagatcccca	gggtcttcac	ctcctctatt	aagacaaaat	ctgcagaaat	gacgatcacc
	16021	acccaagcaa	gtcctcctga	gtctgcacgc	cacagtaccc	ttcccttgga	cacatcaacc
	16081	acactttccc	agggagggac	tcattcaact	gtgactcagg	gattcccata	ctcagagggtg
	16141	accactctca	tgggcatggg	tcttggaat	gtgtcatgga	tgacaactcc	cctgtggga
45	16201	gaaaccagct	ctgtgtcttc	cctgatgtct	tcacctgcca	tgacatcccc	ttctcctggt
	16261	tcctccacat	caccacagag	catccccctc	tctcctcttc	ctgtgactgc	acttcctact
	16321	tctgttctgg	tgacaaccac	agatgtgttg	ggcacaacaa	gcccagagtc	tgtaaccagt
	16381	tcacctccaa	atttgagcag	catcactcat	gagagaccgg	ccacttacaa	agacactgca
	16441	cacacagaag	ccgccatgca	tcattccaca	aacaccgcag	tgaccaatgt	agggaacttc
50	16501	gggtctggac	ataaatcaca	atcctctgtc	ctagctgact	cagagacatc	gaaagccaca
	16561	cctctgatga	gtaccacctc	cacctggggg	gacacaaagt	tttccacatc	aactccta
	16621	atctctcaga	ctaaccacaa	tcaaacagag	ccaacagcat	ccttgagccc	tagactgagg
	16681	gagagcagca	cgtctgagaa	gaccagctca	acaacagaga	caaatactgc	cttttcttat
	16741	gtgcccacag	gtgctattac	tcaggcctcc	agaacagaaa	tctcctctag	cagaacatcc
55	16801	atctcagacc	ttgatcggcc	cacaatagca	cccgacatct	ccacaggaat	gatcaccagg
	16861	ctcttcacct	cccccatcat	gacaaaatct	gcagaaaatga	cogtcaccac	tcaaacacact
	16921	actcctgggg	ctacatcaca	gggtatcctt	ccttggggaca	catcaaccac	acttttccag
	16981	ggagggactc	attcaaccgt	gtctcaggga	ttcccacact	cagagataac	cactcttcgg
	17041	agcagaaccc	ctggagatgt	gtcatggatg	acaactcccc	ctgtggaaga	aaccagctct
60	17101	gggttttccc	tgatgtcacc	ttccatgaca	tcccctcttc	ctgtttctct	cacatoccca
	17161	gagagcatcc	cctcctctcc	tctcctctgt	actgcacttc	ttacttctgt	tctgggtgaca

	17221	accaccaatg	tattggggcac	aacaagccca	gagacogtaa	cgagttcacc	tccaaattta
	17281	agcagcccca	cacaggagag	actgaccact	tacaaagaca	ctgcgcacac	agaagccatg
	17341	catgcttcca	tgcatacaaa	cactgcagtg	gccaacgtcg	ggacctccat	ttctggacat
5	17401	gaatcacaa	cttctgtccc	agctgattca	cacacatcca	aagccacatc	tccaatgggt
	17461	atcaccttcg	ccatggggga	tacaagtgtt	tctacatcaa	ctcctgcctt	ctttgagact
	17521	agaattcaga	ctgaatcaac	atcctctttg	attcctggat	taagggacac	caggacgtct
	17581	gaggagatca	acactgtgac	agagaccagc	actgtccttt	cagaagtggc	cactactact
	17641	actactgagg	tctccaggac	agaagttatc	acttccagca	gaacaacccat	ctcagggcct
10	17701	gatcattcca	aaatgtcacc	ctacatctcc	acagaaaacca	tcaccaggct	ctccactttt
	17761	ccttttgtaa	caggatccac	agaaatggcc	atcaccaacc	aaacagggtcc	tatagggact
	17821	atctcacagg	ctaccctttac	cctggacaca	tcaagcacag	cttctgggga	agggactcac
	17881	tcacctgtga	ctcagagatt	tccacactca	gaggagacca	ctactatgag	cagaagtact
	17941	aagggcggtg	catggcagaag	ccctccctct	gtgggaagaaa	ccagttctcc	ttcttcccca
15	18001	gtgcctttac	ctgcaataac	ctcacattca	tctctttatt	ccgcagtatc	aggaagtatg
	18061	cccacttctg	ctctccctgt	gacttccctt	ctcacctctg	gcaggaggaa	gaccatagac
	18121	atgttgagaca	cacactcaga	acttggtgacc	agctccttac	caagtgcgaag	tagcttctca
	18181	ggtgagatac	tcaacttctga	agcctccaca	aatacagaga	caattcactt	ttcagagAAC
	18241	acagcagaaa	ccaatatggg	gaccaccaat	tctatgcata	aactacattc	ctctgtctca
20	18301	atccactccc	agccatccgg	acacacacct	ccaaagggtta	ctggatctat	gatggaggac
	18361	gctattgttt	ccacatcaac	acctgggtct	cctgagacta	aaaatgttga	cagagactca
	18421	acatccccct	tgactcctga	actgaaagag	gacagcaccc	ccctgggtgat	gaactcaact
	18481	acagagtcaa	acactgtttt	ctccagtgtg	tccttggtatg	ctgctactga	ggctctccagg
	18541	cgagaagtca	cctactatga	tcctacattc	atgccagctt	ctgctcagtc	aacaaagtcc
25	18601	ccagacattt	cacctgaagc	cagcagcagt	cattctaaact	ctctccctt	gacaatatct
	18661	acacacaaga	ccatcgccac	acaaacaggt	ccttctgggg	tgacatctct	tggccaactg
	18721	accctggaca	catcaacccat	agccacctca	gcaggaaactc	catcagccag	aactcaggat
	18781	tttgtagatt	cagaaacaac	cagtgtcatg	aacaatgatc	tcaatgatgt	gttgaagaca
	18841	agccctttct	ctgcagaaga	agccaaactc	ctctcttctc	aggcacctct	ccttgtgaca
30	18901	acctcacctt	ctcctgtaac	ttccacattg	caagagcaca	gtacctctc	cttctgttct
	18961	gtgacctcag	taccaccccc	tacactggcg	aagatcacag	acatggacac	aaacttagaa
	19021	cctgtgactc	gttcacctca	aaatttaagg	aacaccttgg	ccacttcaga	agccaccaca
	19081	gatacacaca	caatgcatcc	ttctataaac	acagcaatgg	ccaatgtggg	gaccaccagt
	19141	tcaccaaatg	aattctattt	tactgtctca	cctgactcag	acccatataa	agccacatcc
35	19201	gcagtagtta	tcaacttcac	ctcgggggac	tcaatagtgt	ccacatcaat	gcctagatcc
	19261	tctgcgatga	aaaagattga	gtctgagaca	actttctccc	tgatatttag	actgagggag
	19321	actagcacct	cccagaaaat	tgggtcatcc	tcagacacaa	gcacggtctt	tgacaaagca
	19381	ttcactgctg	ctactactga	ggctctccaga	acagaactca	cctcctctag	cagaacatcc
	19441	atccaaggca	ctgaaaagcc	cacaatgtca	ccggacacct	ccacaagatc	tgtcaccatg
40	19501	ctttctactt	ttgctggcct	gacaaaatcc	gaagaaagga	ccattggcac	ccaaacaggt
	19561	cctcataggg	cgacatcaca	gggtaccctt	acctgggaca	catcaatcac	aacctcacag
	19621	gcagggaccc	actcagctat	gactcatgga	ttttcacaa	tagatttgtc	cactcttacg
	19681	agtagagttc	ctgagtagat	atcagggaca	agccccacct	ctgtggaaaa	aaccagctct
	19741	tcctcttccc	ttctgtcttt	accagcaata	acctcaccgt	cccctgtacc	tactacatta
45	19801	ccagaaagta	ggcgtctctc	tcctgttcat	ctgacttcac	tccccacctc	tggcctagtg
	19861	aagaccacag	atatgctggc	atctgtggcc	agtttacctc	caaacttggg	cagcacctca
	19921	cataagatac	cgactacttc	agaagacatt	aaagatacag	agaaaatgta	tccttcacac
	19981	aacatagcag	taaccaatgt	ggggaccacc	acttctgaaa	aggaatctta	ttcgtctgtc
50	20041	ccagcctact	cagaaccacc	caaagtccacc	tctccaatgg	ttacctcttt	caacataagg
	20101	gacaccattg	tttccacatc	catgcctggc	tcctctgaga	ttacaaggat	tgagtggag
	20161	tcaacattct	ccgtggctca	tgggtggaag	ggaaccagca	cctcccagga	ccccatcgta
	20221	tccacagaga	aaagtgtgtg	ccttcacaag	ttgaccactg	gtgctactga	gacctctagg
	20281	acagaagttg	cctcttctag	aagaacatcc	attccaggcc	ctgatcatcc	cacagagtca
	20341	ccagacatct	ccactgaagt	gatccccagc	ctgacctatc	cccttggcat	tacagaatct
55	20401	tcaaataatga	ccatcatcac	tcgaacaggt	cctcctcttg	gctctacatc	acagggcaca
	20461	tttaccttgg	acacaccaac	tacatcctcc	agggcaggaa	cacactcgat	ggcgactcag
	20521	gaattttccac	actcagaaat	gaccactgtc	atgaacaagg	acccctgagat	tctatcatgg
	20581	acaatccctc	cttctataga	gaaaaccagc	ttctcctctt	ccctgatgcc	ttcaccagcc
	20641	atgacttcac	ctcctgtttc	ctcaacatta	ccaaagacca	ttcacaccac	tccttctcct
60	20701	atgacctcac	tgctcacccc	tagcctagtgt	atgaccacag	acacattggg	cacaagccca
	20761	gaacctacaa	ccagttcacc	tccaaatttg	agcagtacct	cacatgtgat	actgacaaca
	20821	gatgaagaca	ccacagctat	agaagccatg	catccttcca	caagcacagc	agcgactaat

	20881	gtggaaacca	cctgttctgg	acatgggtca	caatcctctg	tcctaactga	ctcagaaaaa
	20941	accaaggcca	cagctccaat	ggataccacc	tcaccatggg	ggcatacaac	tggtttccaca
	21001	tcaatgtctg	tttctcttga	gactacaaaa	attaagagag	agtcaacata	ttccttgact
5	21061	cctggactga	gagagaccag	catttcccaa	aatgccagct	tttccactga	cacaagtatt
	21121	gttctttcag	aagtccccac	tggtactact	gctgaggtct	ccaggacaga	agtcacctcc
	21181	tctggtagaa	catccatccc	tggcccttct	cagtccacag	ttttgccaga	aatatccaca
	21241	agaacaatga	caaggctctt	tgcctcgccc	accatgacag	aatcagcaga	aatgaccatc
	21301	cccactcaaa	caggtccttc	tgggtctacc	tcacaggata	cccttaoctt	ggacacatcc
	21361	accacaaagt	cccaggcaaa	gactcattca	actttgactc	agagatttcc	acactcagag
10	21421	atgaccactc	tcctgagcag	aggctcctgga	gatatgtcat	ggcaaagctc	tcctctctctg
	21481	gaaaatccca	gctctctccc	ttccctgctg	tctttacctg	ccacaacctc	acctcctccc
	21541	atttctctca	cattaccagt	gactatctcc	tcctctctcc	ttcctgtgac	ttcacttctc
	21601	acctctagcc	cggtaacgac	cacagacatg	ttacacacaa	gcccaagact	tgtaaccagt
	21661	tcacctccaa	agctgagcca	cacttcagat	gagagactga	ccactggcaa	ggacaccaca
15	21721	aatacagaag	ctgtgcatcc	ttccacaaac	acagcagcgt	ccaatgtgga	gattcccagc
	21781	tttggacatg	aatccccttc	ctctgacctt	gctgactcag	agacatccaa	agccacatca
	21841	ccaatgttta	ttacctccac	ccaggaggat	acaactgttg	ccatatcaac	ccctcacttc
	21901	ttggagacta	gcagaattca	gaaagagtca	atttctctcc	tgagccctaa	attgaggagg
20	21961	acaggcagtt	ctgtggagac	aagctcagcc	atagagacaa	gtgctgtcct	ttctgaagtg
	22021	tcatttggtg	ctactactga	gatctccagg	acagaagtca	cctcctctag	cagaacatcc
	22081	atctctgggt	ctgctgagtc	cacaatgttg	ccagaaatat	ccaccacaag	aaaaatcatt
	22141	aagttcccta	cttcccccat	cctggcgaaa	tcacagaaaa	tgaccatcaa	gacccaaaca
	22201	agtcctcctg	ggtctacatc	agagagtacc	tttacattag	acacatcaac	cactccctcc
25	22261	ttggaataaa	cccattcgac	tatgactcag	agattgccac	actcagagat	aacctctott
	22321	gtgagtagag	gtgctgggga	tgtgccacgg	cccagctctc	tcctgttgga	agaaacaagc
	22381	cctccatctt	cccagctgtc	tttatctgcc	atgatctcac	cttctcctgt	ttcttcacaa
	22441	ttaccagcaa	gtagccactc	ctcttctgct	tctgtgactt	cacctctcac	accaggccaa
	22501	gtgaagacta	ctgagggtgt	ggacgcgaagt	gcagaacctg	aaaccagttc	acctccaagt
30	22561	ttgagcagca	cctcagttga	aatactggcc	acctctgaag	tcaccacaga	tacggagaaa
	22621	attcatcctt	ttccaaacac	ggcagtaacc	aaagttggaa	cttcacagttc	tggaacatgaa
	22681	ttcccttcct	ctgtcctacc	tgactcagag	acaaccaaag	ccacatcggc	aatgggtacc
	22741	atctccatta	tgggggatac	aagtgtttct	acattaaact	ctgccttata	taacactagg
	22801	aaaattcagt	cagagccagc	ttcctcactg	accaccagat	tgaggagagc	cagcacctct
35	22861	gaagagacca	gcttagccac	agaagcaaac	actgttcttt	ctaaagtgtc	cactggtgct
	22921	actactgagg	tctccaggac	agaagccatc	tccttttagca	gaacatccat	gtcaggccct
	22981	gagcagtcga	caatgtcaca	agacatctcc	ataggaacca	ttcccaggat	ttctgcctcc
	23041	tctgtcctga	cagaatctgc	aaaaatgacc	atcacaaacc	aaacagggtcc	ttcggagtct
40	23101	acactagaaa	gtacccttaa	tttgaacaca	gcaaccacac	cctcttgggt	ggaaacccac
	23161	tctatagtaa	ttcagggtatt	ttcacaccca	gagatgacca	cttccatggg	cagaggctct
	23221	ggagggtgtg	catggcctag	ccctcccttt	gtgaaagaaa	ccagccctcc	atcctccccg
	23281	ctgtctttac	ctgccgtgac	ctcacctcat	cctgtttcca	ccacattcct	agcacatatc
	23341	ccccctctc	cccttctgt	gacttcaett	ctcacctctg	gcccggcgac	aaccacagat
	23401	atcttgggta	caagcacaga	acctggaacc	agttcatctt	caagtttgag	caccacctcc
45	23461	catgagagac	tgaccactta	caaagacact	gcacatacag	aagccgtgca	tccttcacaa
	23521	aacacaggag	ggaccaatgt	ggcaaccacc	agctctggat	ataaatcaca	gtcctctgtc
	23581	ctagctgact	catctccaat	gtgtaccacc	tcaccatggg	gggatacaag	tggtctcaca
	23641	tcaactcctg	ccttctctga	gactaggagg	attcagacag	agctagcttc	ctccttgacc
	23701	cctggattga	gggagtcag	tggctctgaa	gggaccagct	caggacccaa	gatgagcact
50	23761	gtcctctcta	aagtgcccac	tgggtctact	actgagatct	ccaaggaaga	cgtcacctcc
	23821	atcccaggtc	ccgctcaatc	cacaatatca	ccagacatct	ccacaagaac	cgtcagctgg
	23881	ttctctacat	cccctgtcat	gacagaatca	gcagaaataa	ccatgaacac	ccatacaagt
	23941	cctttagggg	ccacaacaca	aggcaccagt	actttggcca	cgtcaagcac	aacctctttg
	24001	acaatgacac	actcaactat	atctcaagga	ttttcacact	cacagatgag	cactcttatg
55	24061	aggaggggtc	ctgaggatgt	atcatggatg	agccctcccc	ttctggaaaa	aactagacct
	24121	tccttttctc	tgatgtcttc	accagccaca	acttcacctt	ctcctgtttc	ctccacatta
	24181	ccagagagca	tctcttcttc	tcctcttctc	gtgacttcac	tcctcacgtc	tggttgggca
	24241	aaaactacag	atatgttgca	caaaagctca	gaacctgtaa	ccaactcacc	tgcaaatattg
	24301	agcagcacct	cagttgaaat	actggccacc	tctgaagtca	ccacagatac	agagaaaact
	24361	catccttctt	caaacagaac	agtgaccgat	gtggggacct	ccagttctgg	acatgaatcc
60	24421	acttctcttg	tcctagctga	ctcacagaca	tcacaaagtca	catctccaat	ggttattacc
	24481	tcaccatggg	aggatacag	tgtctccaca	tcaactcctg	gcttttttga	gactagcaga

	24541	attcagacag	aaccaacatc	ctccctgacc	cttgactga	gaaagaccag	cagctctgag
	24601	gggaccagct	tagccacaga	gatgagcact	gtcctttctg	gagtgcccac	tggtgccact
	24661	gctgaagtct	ccaggacaga	agtcacctcc	tctagcagaa	catccatctc	aggctttgct
	24721	cagctcacag	tgtcaccaga	gacttccaca	gaaacccatca	ccagactccc	tacctccagc
5	24781	ataatgacag	aatcagcaga	aatgatgatc	aagacacaaa	cagatcctcc	tgggtctaca
	24841	ccagagagta	ctcatactgt	ggacatatca	acaacaccca	actgggtaga	aacccactcg
	24901	actgtgactc	agagattttc	acactcagag	atgaccactc	ttgtgagcag	aagccctggg
	24961	gatatgttat	ggcctagtca	atcctctgtg	gaagaaaacca	gctctgcctc	ttccctgctg
	25021	tctctgcctg	ccacgacctc	accttctcct	gtttcctcta	cattagttaga	ggatttccct
10	25081	tccgcttctc	ttcctgtgac	ttctcttctc	acccctggcc	tggtgataac	cacagacagg
	25141	atgggcataa	gcagagaacc	tggaaaccag	tccacttcaa	atttgagcag	cacctcccat
	25201	gagagactga	ccactttgga	agacactgta	gatacagaag	acatgcagcc	ttccacacac
	25261	acagcagtga	ccaacgtgag	gacctccatt	tctggacatg	aatcacaaac	ttctgtccta
	25321	tctgactcag	agacacccaa	agccacatct	ccaatgggta	ccacctacac	catgggggaa
15	25381	acgagtgttt	ccatatccac	ttctgacttc	tttgagacca	gcagaattca	gatagaacca
	25441	acatcctccc	tgacttctgg	attgagggag	accagcagct	ctgagaggat	cagctcagcc
	25501	acagagggaa	gcactgtcct	ttctgaagtg	cccagtgggtg	ctaccactga	ggtctccagg
	25561	acagaagtga	tatcctctag	gggaacatcc	atgtcagggc	ctgatcagtt	caccatatca
	25621	ccagacatct	ctactgaagc	gatccaccag	cttttctact	ccccatttat	gacagaacca
20	25681	gcagaaagtg	ccatcactat	tgagacaggt	tctcctgggg	ctacatcaga	gggtaccctc
	25741	accttggaac	cctcaacaac	aaccttttgg	tcagggaccc	actcaactgc	atctccagga
	25801	ttttcacact	cagagatgac	cactcttatg	agtagaactc	ctggagatgt	gccatggccg
	25861	agccttccct	ctgtggaaga	agccagctct	gtctcttctc	cactgtcttc	acctgccatg
	25921	acctcaactt	cttttttctc	cgcattacca	gagagcatct	cctcctctcc	tcactctgtg
25	25981	actgcacttc	tcacccttgg	cccagtgaag	accacagaca	tgttgccgac	aagctcagaa
	26041	cctgaaacca	gttcacctcc	aaatttgagc	agcacctcag	ctgaaatatt	agccacgtct
	26101	gaagtcacca	aagatagaga	gaaaatttcat	ccctcctcaa	acacacctgt	agtcaatgta
	26161	gggactgtga	tttataaaca	tctatccctc	tcctctgttt	tggctgactt	agtgacaaca
	26221	aaaccccat	ctccaatggc	taccacctcc	actctgggga	atacaagttg	ttccacatca
30	26281	actcctgcct	tcacagaaac	tatgatgaca	cagccaactt	cctccctgac	ttctggatta
	26341	agggagatca	gtacctctca	agagaccagc	tcagcaacag	agagaagtgc	ttctctttct
	26401	ggaatgccca	ctggtgctac	tactaaggtc	tcacagaacag	aagccctctc	cttaggcaga
	26461	acatccaccc	caggtcctgc	tcaatccaca	atatcaccag	aaatctccac	ggaaacccatc
	26521	actagaatth	ctactccctc	caccacgaca	ggatcagcag	aaatgacctc	cacccccaaa
35	26581	acaggtcatt	ctggggcctc	ctcacaaagt	acctttacct	tggacacatc	aagcagagcc
	26641	tcctggccag	gaactcactc	agctgcaact	cacagatctc	cacactcagg	gatgaccact
	26701	cctatgagca	gaggtcctga	ggatgtgtca	tggccaagcc	gcccatcagt	ggaaaaaact
	26761	agccctccat	cttccctggg	gtctttatct	gcagtaacct	caccttgcgc	acctttattcc
	26821	acaccatctg	agagtagcca	ctcatctcct	ctccgggtga	cttctctttt	actccctgtc
40	26881	atgatgaaga	ccacagacat	gttgacacac	agcttggaac	ctgtgaccac	ttcacctccc
	26941	agtatgaata	tcacctcaga	tgagagtctg	gccacttcta	aagccacctc	ggagacagag
	27001	gcaattcagc	tttcagaaaa	cacagctgtg	actcagatgg	gcaccatcag	cgctagacaa
	27061	gaattctatt	cctcttatcc	aggcctccca	gagccatcca	aagtgcacac	tcacagtggtc
	27121	acctcttcca	ccataaaaaga	cattgtttct	acaaccatac	ctgcttctct	tgagataaca
45	27181	agaattgaga	tggagtcaac	atccaccctg	acccccacac	caagggagac	cagcacctcc
	27241	caggagatcc	actcagccac	aaagccaagc	actgttccct	acaaggcact	cactagtggc
	27301	acgattgagg	actccatgac	acaagtcatg	tcctctagca	gaggacctag	ccctgatcag
	27361	tcacacaatg	cacaagacat	atccagtga	gtgatcacca	ggctctctac	ctccccatc
50	27421	aaggcagaat	ctacagaaat	gaccattacc	acccaaacag	gttctctctg	ggctacatca
	27481	aggggtaccc	ttaccttgga	cacttcaaca	acttttatgt	cagggaccca	ctcaactgca
	27541	tctcaaggat	tttcacactc	acagatgacc	gctcttatga	gtagaactcc	tggagatgtg
	27601	ccatggctaa	gccatccctc	tgtggaagaa	gccagctctg	cctctttctc	actgtcttca
	27661	cctgtcatga	cctcatcttc	tcocgtttct	tcacatttac	cagacagcat	ccactcttct
	27721	tcgcttctctg	tgacatcact	tctcacctca	gggctgggtga	agaccacaga	gctgttgggc
55	27781	acaagctcag	aacctgaaac	cagttcaccc	ccaaatttga	gcagcacctc	agctgaaata
	27841	ctggccacca	ctgaagtcac	tacagataca	gagaaactgg	agatgaccaa	tgtggtaacc
	27901	tcaggttata	catatgaatc	tccttctctc	gtcctagctg	actcagtgc	aacaaaggcc
	27961	acatcttcaa	tgggtatcac	ctacccaca	ggagatacaa	atgttctcac	atcaaccctc
	28021	gccttctctg	acaccagtag	gattcaaaaca	aagtcaaagc	tctcactgac	tctgggttg
60	28081	atggagacca	gcattctctga	agagaccagc	tctgccacag	aaaaaagcac	tgtcctttct
	28141	agtggtgcca	ctggtgctac	tactgaggtc	tcacaggacag	aagccatctc	ttctagcaga

	28201	acatccatcc	caggccctgc	tcaatccaca	atgtcatcag	acacctccat	ggaaaccatc
	28261	actagaatth	ctacccccct	cacaaggaaa	gaatcaacag	acatggccat	cacccccaaa
	28321	acaggctcct	ctggggctac	ctcgcagggt	acctttacct	tggactcatc	aagcacagcc
	28381	tccctggccag	gaactcactc	agctacaact	cagagatttc	cacagtcagt	ggtagacaact
5	28441	cctatgagca	gaggctcctga	ggatgtgtca	tggccaagcc	cgctgtctgt	ggaaaaaaac
	28501	agccctccat	cttccctggt	atcttcatct	tcagtaacct	caccttcgcc	actttattcc
	28561	acaccatctg	ggagtagcca	ctcctctcct	gtccctgtca	cttctctttt	cacctctatc
	28621	atgatgaagg	ccacagacat	gttggatgca	agtttggaac	ctgagaccac	ttcagctccc
	28681	aatatgaata	tcacctcaga	tgagagtctg	gccacttcta	aagccaccac	ggagacagag
10	28741	gcaattcacg	tttttgaaaa	tacagcagcg	tcccatgtgg	aaaccaccag	tgctacagag
	28801	gaactctatt	cctcttcccc	aggcttctca	gagccaacaa	aagtgatata	tccagtggct
	28861	acctcttccct	ctataagaga	caacatgggt	tccacaacaa	tgcttggtct	ctctggcatt
	28921	acaaggattg	agatagagtc	aatgtcatct	ctgacccctg	gactgagggg	gaccagaacc
	28981	tcccaggaca	tcacctcatc	cacagagaca	agcactgtcc	tttacaagat	ttcctctggt
15	29041	gccactcctg	aggctctccag	gacagaagtt	atgccctcta	gcagaacatc	cattcctggc
	29101	cctgtctcag	ccacaatgtc	actagacatc	tccgatgaag	ttgtcaccag	gctgtctacc
	29161	tctcccatca	tgacagaatc	tgcagaaata	accatcacca	cccaaacagg	ttattctctg
	29221	gtacatccc	aggttaccct	tcccttgggc	acctcaatga	cctttttgtc	agggaccac
	29281	tcaactatgt	ctcaaggact	ttcacactca	gagatgacca	atcttatgag	caggggctcc
20	29341	gaaagtctgt	catggacgag	ccctcgcttt	gtggaaacaa	ctagatcttc	ctcttctctg
	29401	acatcattac	ctctcacgac	ctcactttct	cctgtgtcct	ccacattact	agacagtagc
	29461	cctctctctc	ctcttctctg	gacttcactt	atcctcccag	gcctgggtga	gactacagaa
	29521	gtgttgata	caagctcaga	gcctaaaacc	agttcatctc	caaatttgag	cagcacctca
	29581	gttgaaatac	cggccacctc	tgaaatcagt	acagatacag	agaaaattca	tctctctca
25	29641	aacacagcgg	tggccaaagt	gaggacctcc	agttctgttc	atgaatctca	ttcctctgtc
	29701	ctagctgact	cagaaacaac	cataaccata	ccttcaatgg	gtatcacctc	cgctgtggac
	29761	gataccactg	ttttcacatc	aaatcctgcc	ttctctgaga	ctaggaggat	tccgacagag
	29821	ccaacattct	cattgactcc	tggattcagg	gagactagca	cctctgaaga	gaccacctca
	29881	atcacagaaa	caagtgagct	cctttatgga	gtgcccacta	gtgtactact	tgaagtctcc
30	29941	atgacagaaa	tcatgtctct	taatagaaca	cacatccctg	actctgatca	gtccacgatg
	30001	tctccagaca	tcatcaactga	agtgatcacc	aggctctctt	cctcatccat	gatgtcagaa
	30061	tcaacacaaa	tgaccatcac	cacccaaaaa	agttctcctg	gggctacagc	acagagtact
	30121	cttaccttgg	ccacaacaac	agcccccttg	gcaaggaccc	actcaactgt	tcctcctaga
35	30181	tttttactac	cagagatgac	aactcctatg	agtaggagtc	ctgaaaatcc	atcatggaag
	30241	agctctccct	ttgtggaaaa	aactagctct	tcactctctc	tggtgtcctt	acctgtcacg
	30301	acctcacctt	ctgtttcttc	cacattaccg	cagagtatcc	cttctctctc	tttttctgtg
	30361	acttcaactc	tcaccccagg	catgggtgaag	actacagaca	caagcacaga	acctggaacc
	30421	agttttatct	caaatctgag	tggcaccttc	gttgaaatac	tggtgtcctc	tgaagtccac
	30481	acagatacag	agaaaattca	tccttcttca	agcatggcag	tgaccaatgt	gggaaccacc
40	30541	agttctggac	atgaactata	ttcctctgtt	tcaatccact	cggagccatc	caaggctaca
	30601	taccacagtg	gtactccctc	ttccatggct	gaaacctota	tttccacatc	aatgcctgct
	30661	aatttttgaga	ccacaggatt	tgaggctgag	ccattttctc	atttgacttc	tggatttagg
	30721	aagacaaaaca	tgtccctgga	caccagctca	gtcacaccaa	caaatacacc	ttcttctcct
	30781	gggtccactc	acctttttaca	gagttccaag	actgatttca	cctcttctgc	aaaaacatca
45	30841	tcccagact	ggcctccagc	ctcacagtat	actgaaatcc	cagtggacat	aatcaccccc
	30901	tttaagtctt	ctccatctat	tacggagtcc	actgggataa	cctccttccc	agaatccagg
	30961	tttactatgt	ctgtaacaga	aagtactcat	catctgagta	cagatttgct	gccttcagct
	31021	gagactatth	ccactggcac	agtgatgcct	tctctatcag	aggccatgac	ttcatttgcc
	31081	accactggag	ttccacgagc	catctcaggt	tcaggtagtc	cattctctag	gacagagtca
50	31141	ggccctgggg	atgctactct	gtccaccatt	gcagagagcc	tgccctcatc	cactcctgtg
	31201	ccattctcct	cttcaacctt	cactaccact	gattcttcaa	ccatcccagc	cctccatgag
	31261	ataacttccct	cttcagctac	cccatataga	gtggacacca	gtcttgggac	agagagcagc
	31321	actactgaag	gacgcttggg	tatggctcagt	actttggaca	cttcaagcca	accaggcagg
	31381	acatcttcaa	cacccatttt	ggataccaga	atgacagaga	gcgttgagct	gggaacagtg
55	31441	acaagtgtct	atcaagttcc	ttcactctca	acacgggtga	caagaactga	tggcattatg
	31501	gaacacatca	caaaaatacc	caatgaagca	gcacacagag	gtaccataag	accagtcaaa
	31561	ggccctcaga	catccacttc	gcctgccagt	cctaagggac	tacacacagg	agggacaaaa
	31621	agaatggaga	ccaccaccac	agctttgaag	accaccacca	cagctttgaa	gaccttcc
	31681	agagccacct	tgaccaccag	tgtctatact	cccactttgg	gaacactgac	tcccctcaat
60	31741	gcatcaaggc	aaatggccag	cacaatcctc	acagaaatga	tgatcacaa	cccatatgtt
	31801	ttccctgatg	ttccagaaac	gacatcctca	ttggctacca	gcctgggagc	agaaaccagc

	31861	acagctcttc	ccaggacaac	cccatctggt	ctcaatagag	aatcagagac	cacagcctca
	31921	ctgggtctctc	gttctggggc	agagagaagt	cgggttattc	aaactctaga	tgtttcttct
	31981	agtggagccag	atacaacagc	ttcatgggtt	atccatcctg	cagagacccat	cccaactggt
5	32041	tccaagacaa	cccccaattt	ttccacagt	gaattagaca	ctgtatcttc	cacagccacc
	32101	agtcagggg	cagacgtcag	ctcagccatt	ccaacaaata	tctcacctag	tgaactagat
	32161	gcactgaccc	cactggtcac	tatttcgggg	acagatacta	gtacaacatt	cccaactctg
	32221	actaagtcctc	cacatgaaac	agagacaaga	accacatggc	tcaactcatcc	tgcagagacc
	32281	agctcaacta	ttcccagaac	aatccccaat	ttttctcatc	atgaatcaga	tgccacacct
	32341	tcaatagcca	ccagtcctgg	ggcagaaacc	agttcagcta	ttccaattat	gactgtctca
10	32401	cctgggtgcag	aagatctggg	gacctcacag	gtcactagtt	ctgggacaga	cagaaatag
	32461	actattccaa	ctttgactct	ttctcctggg	gaaccaaaga	cgatagcctc	attagtcacc
	32521	catcctgaag	cacagacaag	ttcggccatt	ccaacttcaa	ctatctcgcc	tgctgtatca
	32581	cggttggtga	cctcaatggg	caccagtttg	gcggaacaaga	caagtacaac	taatcgagct
	32641	ctgacaaact	ccctgggtga	accagctaca	acagtttcat	tggtcacgca	tccgtgcacag
15	32701	accagcccaa	cagttccctg	gacaacttcc	atTTTTTTTcc	atagtaaatac	agacaccaca
	32761	ccttcaatga	ccaccagtc	tggggcagaa	tccagttcag	ctgttccaac	tccaactggt
	32821	tcaactgagg	taccaggagt	agtgacctct	ttgggtcacca	gttctagggc	agtgatcagt
	32881	acaactatcc	caattctgac	tctttctcct	gggtgaaccag	agaccacacc	ttcaatggcc
	32941	accagtcctg	gggaagaagc	cagttctgct	attccaactc	caactgtttc	acctggggt
20	33001	ccaggagtgg	tgacctctct	gggtcactagt	tctagggcag	tgactagtag	aactattcca
	33061	attctgactt	tttctcttgg	tgaaccagag	accacacctt	caatggccac	cagtcagggg
	33121	acagaagctg	gctcagctgt	tccaactggt	ttacctgagg	taccaggaat	gggtgacctct
	33181	ctgggtgcta	gttctagggc	agtaaccagt	acaactcttc	caactctgac	tctttctcct
25	33241	gggtgaaccag	agaccacacc	ttcaatggcc	accagtcag	gggcagaaag	tgactcaact
	33301	gttccaactg	tttcaactga	gggtaccagga	gtgggtgacct	ctctgggtcac	tagttctagt
	33361	ggagtaaaaca	gtacaagtat	tccaactctg	attctttctc	ctgggtgaact	agaaaccaca
	33421	ccttcaatgg	ccaccagtc	tggggcagaa	gccagctcag	ctgttccaac	tccaactggt
	33481	tccctggggg	tatcaggagt	gggtgacctc	ctgggtcact	gttccagggc	agtgaccagt
	33541	acaactatcc	caattctaac	tctttcttct	agtgagccag	agaccacacc	ttcaatggcc
30	33601	accagtcag	gggtagaagc	cagctcagct	gttctaaactg	tttcaactga	gggtaccagga
	33661	atgggtgacct	ctctgggtcac	tagttctaga	gcagtaacca	gtacaactat	tccaactctg
	33721	actatttctt	ctgatgaacc	agagaccaca	acttcattgg	tcaacctatc	tgaggcaaac
	33781	atgatttcag	ccattccaac	tttagctgtc	tccctactg	tacaagggtc	gggtgacttca
35	33841	ctgggtcacta	gttctgggtc	agagaccagt	gcgttttcaa	atctaactgt	tgccctcaagt
	33901	caaccagaga	ccatagactc	atgggtcgct	catcctggga	cagaagcaag	ttctgttggt
	33961	ccaactttga	ctgtctccac	tgggtgagccg	tttacaataa	tctcaattgg	cacctatcct
	34021	gcagagagta	gctcaactct	tcccaggaca	acctcaagggt	tttcccacag	tgaattagac
	34081	actatgcctt	ctacagtcac	cagtcctgag	gcagaatcca	gtcagccat	ttcaactact
40	34141	atcttcaactg	gtataaccag	tgtgctgaca	tcaactgggtc	ctagctctgg	gagagacatc
	34201	agtgcaactt	ttccaacagt	gacctgagtc	ccacatgaat	cagaggcaac	agcctcatgg
	34261	gttactcatc	ctgcagtcac	cagcacaca	gttcccagga	caacctctaa	ttattctcat
	34321	agtgaaccag	acaccacacc	atcaatagcc	accagtcctg	gggcagaagc	cacttcagat
	34381	tttccaacaa	taactgtctc	acctgatgta	ccagatattg	taacctcaca	gggtcactagt
45	34441	tctgggacag	acaccagtat	aactattcca	actctgactc	tttcttctgg	tgagccagag
	34501	accacaacct	catttatcac	ctattctgag	acacacacaa	gttcagccat	tccaactctc
	34561	cctgtctccc	ctgggtgcac	aaagatgctg	acctcactgg	tcatcagttc	tgggacagac
	34621	agcactacaa	ctttcccaac	actgacggag	accccatatg	aaccagagac	aacagccata
	34681	cagctcattc	atcctgcaga	gaccaacaca	atgggttcca	agacaactcc	caagttttcc
50	34741	catagtaagt	cagacaccac	actcccagta	gccatcacca	gtcctggggtc	agaagccagt
	34801	tcagctgttt	caacgacaac	tatctcaact	gatagtctag	atctgggtgac	ctcactgggtc
	34861	cctagttctg	ggacagacac	cagtacaacc	ttcccaacat	tgagttagac	cccatatgaa
	34921	ccagagacta	cagtcacgtg	gctcactcat	cctgcagaaa	ccagcacaa	gggtttctggg
	34981	acaattccca	acttttccca	taggggatca	gacactgcac	cctcaatggg	caccagtcct
55	35041	ggagtagaca	cgagggtcagg	tggttccaact	acaacctatc	cacctagtat	accaggggta
	35101	gtgacctcac	agggtcactag	ttctgcaaca	gacactagta	cagctattcc	aactttgact
	35161	ccttctcctg	gtgaaccaga	gaccacagcc	tcatcagcta	cccatcctgg	gacacagact
	35221	ggcttcaactg	ttccaattcg	gactgttccc	tctagttagc	cagatacaat	ggcttctctgg
	35281	gtcactcatc	ctccacagac	cagcacacct	gtttccagaa	caacctccag	tttttcccat
	35341	agtagtccag	atgccacacc	tgtaattggcc	accagtccta	ggacagaagc	cagttcagct
60	35401	gtactgacaa	caatctcacc	tggtgcacca	gagatgggtga	cttcacagat	cactagttct
	35461	ggggcagcaa	ccagtacaac	tggttccaact	ttgactcatt	ctcctggtat	gccagagacc

	35521	acagccttat	tgagcaccca	tcccagaaca	gggacaagta	aaacatttcc	tgcttcaact
	35581	gtgtttcctc	aagtatcaga	gaccacagcc	tcactcacca	ttagacctgg	tgcagagact
	35641	agcagagctc	tcccaactca	gacaacatcc	tctctcttca	ccctacttgt	aactggaacc
	35701	agcagagtgg	atctaagtc	aactgcttca	cctgggtgtt	ctgcaaaaac	agccccactt
5	35761	tccacccatc	cagggacaga	gaccagcaca	atgattccaa	cttcaactct	ttcccttggg
	35821	ttactagaga	ctacaggctt	actggccacc	agctcttcag	cagagaccag	cacgagtact
	35881	ctaactctga	ctgtttcccc	tgctgtctct	gggctttcca	gtgcctctat	aacaactgat
	35941	aagccccaaa	ctgtgacctc	ctggaacaca	gaaacctcac	catctgtaac	ttcagttgga
	36001	cccccagaat	tttccaggac	tgtoacaggc	accactatga	ccttgatacc	atcagagatg
10	36061	ccaacaccac	ctaaaaccag	tcattggagaa	ggagtggatc	caaccactat	cttgagaact
	36121	acaatgggtg	aagccactaa	tttagctacc	acaggttcca	gtcccaactg	ggccaagaca
	36181	acaaccacct	tcaatacact	ggctggaagc	ctctttactc	ctctgaccac	acctgggatg
	36241	tcacacctgg	cctctgagag	tgtagacctc	agaacaagtt	ataaaccatg	ctcctggatc
	36301	tccaccacca	gcagttataa	ccgtcggtac	tggacctctg	ccaccagcac	tccagtgtact
15	36361	tctacattct	ccccagggat	ttccacatcc	tccatcccca	gctccacagc	agccacagtc
	36421	ccattcatgg	tgccattcac	cctcaacttc	accatcacca	acctgcagta	cgaggaggac
	36481	atgcggcacc	ctggttccag	gaagttcaac	gccacagaga	gagaactgca	gggtctgctc
	36541	aaaccttgt	tcaggaaatag	cagtctggaa	tacctctatt	caggctgcag	actagcctca
	36601	ctcaggccag	agaaggatag	ctcagccatg	gcagtggatg	ccatctgcac	acatcgccct
20	36661	gacctgaag	acctcggaact	ggacagagag	cgactgtact	gggagctgag	caatctgaca
	36721	aatggcatcc	aggagctggg	ccctacacc	ctggaccgga	acagtctcta	tgtcaatggg
	36781	ttcacccatc	gaagctctat	gcccaccacc	agcactcctg	ggacctccac	agtggatgtg
	36841	ggaaacctcag	ggactccatc	ctocagcccc	agccccacgg	ctgctggccc	tctctgtatg
	36901	ccgttcaccc	tcaacttcac	catcaccaac	ctgcagtaac	aggaggacat	gcgtcgact
25	36961	ggctccagga	agttcaacac	catggagagt	gtcctgcagg	gtctgtctca	gcccttggtc
	37021	aagaacacca	gtgttggccc	tctgtactct	ggctgcagat	tgaccttgct	caggccccag
	37081	aaagatgggg	cagccactgg	agtggatgcc	atctgcaccc	accgccttga	ccccaaaagc
	37141	cctggactca	acagggagca	gctgtactgc	gagctaagca	aactgaccaa	tgacattgaa
	37201	gagctggggc	cctacaccct	ggacagggaac	agtctctatg	tcaatgggtt	caccatcag
30	37261	agctctgtgt	ccaccaccag	cactcctggg	acctccacag	tggatctcag	aacctcaggg
	37321	actccatcct	ccctctccag	ccccacaatt	atggctgctg	gcctctcct	ggtaccattc
	37381	acctcaact	tcaccatcac	caacctgcag	tatggggagg	acatgggtca	ccctggctcc
	37441	aggaagtcca	acaccacaga	gagggtcctg	cagggtctgc	ttgggtccat	attcaagaac
	37501	accagtgttg	gcccctctgta	ctctggctgc	agactgacct	ctctcaggtc	tgagaaggat
35	37561	ggagcagcca	ctggagtggg	tgccatctgc	atccatcatc	ttgaccccaa	aagccctgga
	37621	ctcaacagag	agcggctgta	ctgggagctg	agccaaactga	ccaatggcat	caaagagctg
	37681	ggccccatac	ccctggacag	gaacagtctc	tatgtcaatg	gtttcaccca	tcggacctct
	37741	gtgcccccca	ccagcactcc	tgggacctcc	acagtggacc	ttggaaacct	agggactcca
	37801	ttctccctcc	caagccccgc	aaactgctgg	cctctcctgg	tgctgttcac	cctcaacttc
40	37861	accatcacca	acctgaagta	tgaggaggac	atgcatcgcc	ctggctccag	gaagttcaac
	37921	accactgaga	gggtcctgca	gactctgctt	ggtcctatgt	tcaagaacac	cagtgttggc
	37981	cttctgtact	ctggctgcag	actgaccttg	ctcaggtccg	agaaggatgg	agcagccact
	38041	ggagtggatg	ccatctgcac	ccaccgtctt	gacccccaaa	gccctggact	ggacagagag
	38101	cagctatact	gggagctgag	ccagctgacc	aatggcatca	aagagctggg	cccctacacc
45	38161	ctggacagga	acagtctcta	tgtcaatggg	ttcacccatt	ggatccctgt	gcccaccagc
	38221	agcactcctg	ggacctccac	agtggacctt	gggtcagggg	ctccatccct	cctccccagc
	38281	cccacagctg	ctggccctct	cctgggtgcca	ttcacccctc	acttcaccat	caccaacctg
	38341	cagtagcagg	aggacatgca	tcacccaggc	tcagggaagt	tcaacaccac	ggagcgggtc
	38401	ctgcagggtc	tgcttgggtc	catgttcaag	aacaccagtg	tcggccttct	gtactctggc
50	38461	tgcagactga	ccttgtctcag	gtccgagaag	gatggagcag	ccactggagt	ggatgccatc
	38521	tgcacccacc	gtcttgaccc	caaaagccct	ggagtggaca	gggagcagct	atactgggag
	38581	ctgagccagc	tgaccaatgg	catcaaagag	ctgggtccct	acacccctga	cagaaacagt
	38641	ctctatgtca	atgggtttcac	ccatcagacc	tctgcgcccc	acaccagcac	tctggggacc
	38701	tccacagtgg	accttggggc	ctcagggact	ccatcctccc	tccccagccc	tacatcngct
55	38761	ggccctctcc	tggtnccttt	caccctcaac	ttcacccatc	ccaacctgca	gtacgaggag
	38821	gacatgcggc	accnngnntc	caggaagttc	aacaccacng	agagggtnct	gcagggctctg
	38881	ctnaagcccc	tnntcaagag	caccagtgtt	ggccctctgt	actctggctg	cagactgacc
	38941	ttgctcaggt	ccgagaagga	tggagcagcc	actggagtgg	atgccatctg	caccaccgtg
	39001	cttgaccccca	aaagccctgg	agtggacagg	gagcagctat	actgggagct	gagccagctg
60	39061	accaatggca	tcaaagagct	gggtccctac	acctgggaca	gaaacagtct	ctatgtcaat
	39121	ggtttcaccc	atcagacctc	tgcgccccaa	accagcactc	ctgggacctc	cacagtggac

39181 cttgggacct cagggactcc atcctccctc ccagcccta catctgctgg cctctcctg
 39241 gtgccattca ccctcaactt caccatcacc aacctgcagt acgaggagga catgcatcac
 39301 ccaggctcca ggaagtcca caccacggag cgggtcctgc aggtctgctt tgggtcccatg
 39361 tccaagaaca ccagtgtcgg cctctgttac tctggctgca gactgacctt gctcaggcct
 5 39421 gagaagaatg gggcagccac tgggaatggat gccatctgca gccaccgtct tgacccaaa
 39481 agccctggac tcaacagaga gcagctgtac tgggagctga gccagctgac ccatggcatc
 39541 aaagagctgg gcccctacac cctggacagg aacagtctct atgtcaatgg ttccacccat
 39601 cggagctctg tggcccccac cagcactcct gggacctcca cagtggacct tgggacctca
 10 39661 gggactccat cctccctccc cagcccccaca acagctgttc ctctcctggg gcggttcacc
 39721 ctcaacttta ccatcaccaa tctgcagtat ggggaggaca tgcgtcaccg tgggtccagg
 39781 aagttcaaca ccacagagag ggtcctgcag ggtctgcttg gtcccttggt caagaactcc
 39841 agtgtcggcc ctctgtactc tggctgcaga ctgatctctc tcaggctctga gaaggatggg
 39901 gcagccactg gagtggatgc catctgcacc caccacctta accctcaaag ccttgactg
 15 39961 gacagggagc agctgtactg gcagctgagc cagatgacca atggcatcaa agagctgggc
 40021 ccctacacco tggaccgga cagtctctac gtcaatgggt tcccccatcg gagctctggg
 40081 ctcaccacca gactccttg gacttccaca gttgaccttg gaacctcagg gactccatcc
 40141 cccgtcccca gcccacaac tgetggccct ctccctgggc cattcaccct caacttacc
 40201 atcaccaacc tgcagtatga ggaggacatg catcgccctg gatctaggaa gttcaacacc
 20 40261 acagagaggg tctgcaggg tctgcttagt cccattttca agaactccag tgttggccct
 40321 ctgtactctg gctgcagact gacctctctc agggccgaga aggatggggc agcaactgga
 40381 atggatgctg tctgcctcta ccacctaat cccaaaagac ctggactgga cagagagcag
 40441 ctgtactggg agctaagcca gctgaccac aacatcactg agctggggcc ctacagcctg
 40501 gcaggggaca gtctctatgt caatggtttc acccatcaga actctgtgcc caccaccagt
 25 40561 actcctggga cctccacagt gtactgggca accactggga ctccatcctc cttcccgggc
 40621 cacacagagc ctggccctct cctgatacca ttcactttca actttaccat caccaacctg
 40681 cattatgagg aaaacatgca acacctgggt tccaggaagt tcaacaccac ggagaggggt
 40741 ctgcagggtc tgctcaagcc ctgttcaag aacaccagt tggccctct gtactctggc
 40801 tgcagactga cctctctcag gcccgagaag gatggggcag caactggaat ggatgtgtc
 30 40861 tgccctctacc accctaattc caaaagacct gggctggaca gagagcagct ctactgggag
 40921 ctaagccagc tgaccacaa catcactgag ctggggccct acagcctgga cagggacagt
 40981 ctctatgtca atggtttcac ccatacagaac tctgtgcccc ccaccagtac tctgggacc
 41041 tccacagtgt actgggcaac cactgggact ccatacctct tcccggcca cacagagcct
 41101 ggccctctcc tgataccatt cactttcaac ttaccatca ccaactgca ttataggaa
 35 41161 aacatgcaac accctgggtc cagggaagttc aacaccacgg agagggtctc gacgggtctg
 41221 ctcaagccct tgttcaagaa caccagtgtt ggccctctgt actctggctg cagactgacc
 41281 ttgctcagac ctgagaagca tgaggcagcc actggagtg acaccatctg taccaccgc
 41341 gttgatccca tcggacctgg actggacagg gagcggtat actgggagct gagccagctg
 41401 accaaccagc ttaccgaact gggaccctac accctggaca gggacagtct ctatgtcaat
 40 41461 ggcttcaacc ctggagctc tgtgccaacc accagcactc ctgggacctc cagagtgcac
 41521 ctggcaacct ctgggactcc atcctccctg cctggccaca cagccctgt cctctcttg
 41581 ataccattca ccctcaactt taccatcacc aacctgcatt atgaggaaaa catgcaacac
 41641 cctgggtcca ggaagttcaa caccacggag aggggtctgc agggctctgct caagcccttg
 41701 tccaagaaca ccagtgttgg cctctgttac tctggctgca gactgacctt gctcagacct
 45 41761 gagaagcatg aggcagccac tggagtggac accatctgta cccaccgtg tgatccatc
 41821 ggacctggac tgnacagnga gcngctntac tgggagctna gccanctgac caannncatc
 41881 nngagctgg gncctacac cctggacagg nacagtctct atgtcaatgg ttccacccat
 41941 cnganctctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nggnacctcn
 42001 gggactccat cctcctccc cngccncaca tctgctggcc ctctcctggg gccattcacc
 50 42061 ctcaacttca ccatcaccaa cctgcagtac gaggaggaca tgcataccc aggtccagg
 42121 aagttcaaca ccacggagcg ggtcctgcag ggtctgcttg gtccatggt caagaacacc
 42181 agtgtcggcc ttctgtactc tggctgcaga ctgacctgac tcaggcctga gaagaatggg
 42241 gcagccactg gaatggatgc catctgcagc caccgtcttg accccaaaag ccttgactc
 42301 gacagagagc agctgtactg ggagctgagc cagctgacct atggcatcaa agagctgggc
 55 42361 ccctacaccc tggacaggaa cagtctctat gtcaatgggt tcaccatcg gactctgtg
 42421 gccccaccca gactcctgg gacctcaca gtggaccttg ggacctcagg gactccatcc
 42481 tccctcccca gcccacaac agctgttcc ctccctgggc cgttaccct caactttacc
 42541 atcaccaatc tgcagtatgg ggaggacatg cgtcaccctg gctccaggaa gttcaacacc
 42601 acagagaggg tctgcaagg tctgcttggt ccttggttca agaactccag tgtcggccct
 60 42661 ctgtactctg gctgcagact gatctctctc aggtctgaga aggatggggc agccactgga
 42721 gtggatgcca tctgcaccca caaccttaac cctcaaagcc ctggactgga cagggagcag
 42781 ctgtactggc agctgagcca gatgaccaat ggcatacaag agctggggcc ctacaccctg

	42841	gaccggaaca	gtctctacgt	caatgggttc	acccatcgga	gctctgggct	caccaccagc
	42901	actccttggga	cttccacagt	tgaccttggga	acctcagggga	ctccatcccc	cgtccccagc
	42961	cccacaactg	ctggccctct	cctgggtgcca	ttcacccata	acttcacccat	caccaacctg
5	43021	cagtatgagg	aggacatgca	tcgccctgga	tctaggaagt	tcaacgccac	agagagggtc
	43081	ctgcagggtc	tgcttagtcc	catattcaag	aactccagtg	ttggccctct	gtactctggc
	43141	tgcagactga	cctctctcag	gcccagaaag	gatggggcag	caactggaat	ggatgctgtc
	43201	tgccctctacc	accctaatacc	caaaagacct	ggactggaca	gagagcagct	gtactgggag
	43261	ctaagccagc	tgaccacaaa	catcactgag	ctggggccct	acagcctgga	cagggacagt
	43321	ctctatgtca	atgggtttcac	ccatcagagc	tctatgacga	ccaccagaac	tctgatatac
10	43381	tcacacaatgc	acctggcaac	ctcgagaact	ccagcctccc	tgtctggacc	tacgaccgcc
	43441	agccctctcc	tggtgctatt	cacaatcaac	tgcacatca	ccaacctgca	gtacgaggag
	43501	gacatgcgtc	gcactggctc	caggaagtcc	aacacatagg	agagtgtcct	gcagggctctg
	43561	ctcaagccct	tggtcaagaa	caccagtgtt	ggccctctgt	actctggctg	cagattgacc
	43621	ttgctcaggg	ccaagaaaaga	tggggcagcc	actggagtgg	atgccatctg	cacccaccgc
15	43681	cttgacccca	aaagccctgg	actcaacagg	gagcagctgt	actgggagct	aagcaaaactg
	43741	accaatgaca	ttgaagagct	gggcccctac	accctggaca	ggaacagtct	ctatgtcaat
	43801	ggttttcacc	atcagagctc	tgtgtccacc	accagcactc	ctgggacctc	cacagtggat
	43861	ctcagaacct	cagggactcc	atcctccctc	tcagacccca	caattatgnc	nnctgncctt
	43921	ctcctgntnc	cnttccacct	caacttnacc	atcaccaacc	tgcantangn	ggannacatg
20	43981	cnnccccng	gntccaggaa	gttcaacacc	acngagaggg	tcctacaggg	tctgctcagg
	44041	cccttgttca	agaacaccag	tgtcagctct	ctgtactctg	gttgccagact	gaccttgctc
	44101	aggcctgaga	aggatggggc	agccaccaga	gtggatgctg	cctgcacctc	ccgccctgat
	44161	ccc aaaagcc	ctggactgga	cagagagcaa	ctatactggg	agctgagcca	gctaaccac
	44221	agcatcactg	agctggggacc	ctacaccttg	gacaggggtc	gtctctatgt	caatggcttc
25	44281	aaccctcgga	gctctgtgcc	aaccaccagc	actcctggga	cctccacagt	gcacctggca
	44341	acctctggga	ctccatccct	cctgcctggc	cacacancnn	ctgncctctt	cctgntncnn
	44401	ttcacntca	acttnaccat	caccaacctg	cantangngg	annacatgcn	ncnccccngn
	44461	tcagggaagt	tcaacaccac	ngagaggggt	ctgcaggggtc	tgctcaaac	cttgttcagg
	44521	aatagcagtc	tggaataacct	ctattcaggg	tgccagactag	cctcactcag	gccagagaag
30	44581	gatagctcag	ccatggcagt	ggatgccatc	tgcacacatc	gccctgacct	tgaagacctc
	44641	ggactggaca	gagagcgact	gtactgggag	ctgagcaatc	tgcacaaatg	catccaggag
	44701	ctggggccct	acaccctgga	ccggaacagt	ctctacgtca	atgggtttcac	ccatcggagc
	44761	tctgggctca	ccaccagcac	tccttggact	tcacacgttg	accttggaa	ctcagggact
	44821	ccatcccccg	tcaccagccc	cacaactgct	ggccctctcc	tggtgcccatt	cacctcaaac
35	44881	ttcacatca	ccaacctgca	gtatgaggag	gacatgcac	gccctgggtc	caggaggttc
	44941	aacaccacgg	agaggggtct	gcaggggtctg	ctcacgccct	tggtcaagaa	caccagtgtt
	45001	ggccctctgt	actctggctg	cagactgacc	ttgctcagac	ctgagaagca	agaggcagcc
	45061	actggagtgg	acacctctg	taaccaccgc	gttgatocca	tcggacctgg	actggacaga
40	45121	gagcggctat	actgggagct	gagccagctg	accaacagca	tcacagagct	gggaccttac
	45181	accctggata	gggacagtct	ctatgtcaat	ggcttcaacc	cttgagctc	tgtgccaacc
	45241	accagcactc	ctgggacctc	cacagtgcac	ctggcaacct	ctgggactcc	atcctccctg
	45301	cctggccaca	cagccctgtg	ccctctcttg	ataccattca	ccctcaactt	taccatcacc
	45361	gacctgcatt	atgaagaaaa	catgcaacac	cctgggtcca	ggaagttcaa	caccacggag
	45421	aggggtctgc	aggggtctgct	caagcccttg	ttcaagagca	ccagcggtgg	ccctctgtac
45	45481	tctggctgca	gactgacctt	gctcagacct	gagaaacatg	gggcagccac	tggagtggac
	45541	gccatctgca	ccctccgcct	tgatcccaact	ggtcctggac	tggacagaga	gcggctatac
	45601	tgggagctga	gccagctgac	caacagcggt	acagagctgg	gcccctacac	cctggacagg
	45661	gacagtctct	atgtcaatgg	cttcacccat	cggagctctg	tgccaaccac	cagtattcct
	45721	gggacctctg	cagtgcacct	ggaaacctct	gggactccag	cctccctccc	tggccacaca
50	45781	gcccctggcc	ctctcctggg	gccattcacc	ctcaacttca	ctatcaccaa	cctgcagtat
	45841	gaggaggaca	tgctgtcacc	tggttccagg	aagttcagca	ccacggagag	agtcctgcag
	45901	ggtctgtctc	agcccttggt	caagaacacc	agtgtcagct	ctctgtactc	tggttgacga
	45961	ctgaccttgc	tcaggccctga	gaaggatggg	gcagccacca	gagtggatgc	tgtctgcacc
	46021	catcgtctct	accccaaaaag	ccctggactg	gacagagagc	ggctgtactg	gaagctgagc
55	46081	cagctgaccc	acggcatcac	tgagctgggc	ccctacaccc	tggacaggca	cagtctctat
	46141	gtcaatgggt	tcacccatca	gagctctatg	acgaccacca	gaactcctga	tacctccaca
	46201	atgcacctgg	caacctcgag	aactccagcc	tcctgtctgt	gacctacgac	cgccagccct
	46261	ctcctgggtc	tattcaccaat	taacttcacc	atcactaac	tgcggataga	ggagaacatg
	46321	catcacccctg	gctctagaaa	gtttaacacc	acggagagag	tccttcaggg	tctgctcagg
60	46381	cctgtgttca	agaacaccag	tgttggccct	ctgtactctg	gctgcagact	gaccacgctc
	46441	aggcccaaga	aggatggggc	agccaccaaa	gtggatgcca	tctgcacctc	ccgcctgat

	46501	cccaaaagcc	ctggactgga	cagagagcag	ctatactggg	agctgagcca	gctaaccac
	46561	agcatcactg	agctgggccc	ctacaccag	gacagggaca	gtctctatgt	caatggcttc
	46621	acccatcgga	gctctgtgcc	aaccaccagt	attcctggga	cctctgcagt	gcacctggaa
	46681	acctctggga	ctccagcctc	cctccctggc	cacacagccc	ctggccctct	cctgggtcca
5	46741	ttcaccctca	acttcactat	caccaacctg	cagtatgagg	aggacatgcg	tcaccttggt
	46801	tccaggaagt	tcaacaccac	ggagagagtc	ctgcagggtc	tgctcaagcc	cttgttcaag
	46861	agcaccagt	ttggccctct	gtactctggc	tgcagactga	ccttgctcag	gcctgaaaaa
	46921	cgtggggcag	ccaccggcgt	ggacaccatc	tgcactcacc	gccttgaccc	cttaaaccac
	46981	ggactggaca	gagagcagct	atactgggag	ctgagcaaac	tgaccctggg	catcatcgag
10	47041	ctggggccct	acctcctgga	cagaggcagt	ctctatgtca	atggtttcac	ccatcgagcc
	47101	tctgtgcca	ccaccagcac	tctgtggacc	tccacagtgg	accttggaac	ctcagggact
	47161	ccattctccc	tcccaagccc	cgcancnct	gnccctctcc	tgntnccntt	cacnctcaac
	47221	ttnacatca	ccaacctgca	ntangnggan	nacatgcnn	nccnggntc	caggaaagttc
	47281	aacaccacng	agagggtcct	gcagactctg	cttggtccta	tgttcaagaa	caccagtgtt
15	47341	ggccttctgt	actctggctg	cagactgacc	ttgctcaggt	ccgagaagga	tggagcagcc
	47401	actggagtgg	atgccatctg	cacccaccgt	cttgacccca	aaagccctgg	agtggacagg
	47461	gagcaactat	actgggagct	gagccagctg	accaatggca	ttaaagaact	gggcccctac
	47521	acccctggaca	ggaacagctc	ctatgtcaat	gggttcaccc	attggatccc	tgtgcccacc
	47581	agcagcactc	ctgggacctc	cacagtggac	cttggttcag	ggactccatc	ctccctcccc
20	47641	agccccacaa	ctgctggccc	tctcctgggtg	ccgttcaccc	tcaacttcac	catcaccaac
	47701	ctgaagtacg	aggaggacat	gcattgccct	ggctccagga	agttcaacac	cacagagaga
	47761	gtcctgcaga	gtctgcttgg	tcccatgttc	aagaacacca	gtgttggccc	tctgtactct
	47821	gtgtgcagac	tgaccttgct	caggctccag	aaggatggag	cagccactgg	agtggatgcc
	47881	atctgcaccc	accgtcttga	ccccaaaagc	cctggagtgg	acagggagca	gctatactgg
25	47941	gagctgagcc	agctgaccac	tggcatcaaa	gagctgggtc	cctacaccct	ggacagaaac
	48001	agtctctatg	tcaatggttt	cacccatcag	acctctgcgc	ccaacaccag	cactcctggg
	48061	acctccacag	tggaccttgg	gacctcaggg	actccatcct	ccctccccag	ccctacancn
	48121	nctgncccto	tccgtntncc	ntccacntc	aacttnacca	tcaccaacct	gcantangng
	48181	gannacatgc	nnncccnng	ntccaggaag	ttcaacacca	cngagngngt	nctgcagggt
30	48241	ctgctnnnnc	ccntnttcaa	gaacnccagt	gtnggcctc	tgtactctgg	ctgcagactg
	48301	acctnnctca	ggncngagaa	gnatggngca	gccactggan	tggatgccat	ctgcancac
	48361	cnnctnnanc	ccaaaagncc	tggactgnac	agnagcngc	tnactggga	gctnagccan
	48421	ctgaccaann	ncatcnngga	gctgggnccc	tacaccctgg	acaggnacag	tctctatgtc
	48481	aatggtttca	cccatgggat	ccctgtgccc	accagcagca	ctcctgggac	ctccacagtg
35	48541	gaccttgggt	cagggaactcc	atcctccctc	cccagcccca	caactgctgg	ccctctcctg
	48601	gtgccgttca	ccctcaactt	caccatcacc	aacctgaagt	acgaggagga	catgcattgc
	48661	cctggctcca	ggaagttcaa	caccacagag	agagtctgct	agagtctgct	tgttcccctg
	48721	ttcaagaaca	ccagtgttgg	cctctgttac	tctggctgca	gactgacctc	gctcaggtcc
40	48781	gagaaggatg	gagcagccac	tggagtggat	gccatctgca	cccaccgtgt	tgaccccaaa
	48841	agccctggag	tggacaggga	gcagctatac	tgggagctga	gccagctgac	caatggcatc
	48901	aaagagctgg	gtccctacac	cctggacaga	aacagtctct	atgtcaatgg	tttcacccat
	48961	cagacctctg	cgcccaacac	cagcactcct	gggacctcca	cagtgnacnt	nggnacctcn
	49021	gggactccat	cctccntccc	cngccncaca	tctgctggcc	ctctcctggg	gccattccac
	49081	ctcaacttca	ccatcaccaa	cctgcagtac	gaggaggaca	tgcatcacc	aggctccagg
45	49141	aagttcaaca	ccacggagcg	ggtcctgcag	ggtctgcttg	gtcccatggt	caagaacacc
	49201	agtgtcggcc	ttctgtactc	tggctgcaga	ctgaccttgc	tcaggcctga	gaagaatggg
	49261	gcaaccactg	gaatggatgc	catctgcacc	caccgtcttg	accccaaaag	ccctggactg
	49321	nacagnagc	ngctntactg	ggagctnagc	canctgacca	annnccatcn	ngagctgggn
	49381	ccctacaccc	tggacaggna	cagtctctat	gtcaatgggt	tcacccatcn	ganctctgng
50	49441	cccaccacca	gcactcctgg	gacctccaca	gtgnacntng	gnacctcnng	gactccatcc
	49501	tcctntcccn	gccncacanc	nnctgnccct	ctcctgntnc	cnntccacnt	caacttnacc
	49561	atcaccaacc	tgcantangn	ggannacatg	cnnncccnng	gntccaggaa	gttcaacacc
	49621	acngagaggg	ttctgcaggg	tctgctcaaa	cccttggtca	ggaatagcag	tctggaatac
	49681	ctctattcag	gctgcagact	agcctcactc	aggccagaga	aggatagctc	agccatggca
55	49741	gtggatgcca	tctgcacaca	tgcgccctgac	cctgaagacc	tgggactgga	cagagagcga
	49801	ctgtactggg	agctgagcaa	tctgacaaat	ggcatccagg	agctggggcc	ctacaccctg
	49861	gaccggaaca	gtctctatgt	caatgggttc	acccatcgaa	gctctatgct	caccaccagc
	49921	actcctggga	cctccacagt	ggatgtggga	acctcagggg	ctccatcctc	cagccccagc
	49981	cccacgactg	ctggccctct	cctgatacca	ttcaccctca	acttcaccat	caccaacctg
60	50041	cagtatgggg	aggacatggg	tcaccctggc	tccagggaagt	tcaacaccac	agagaggggtc
	50101	ctgcaggggtc	tgcttgggtcc	catattcaag	aacaccagtg	ttggccctct	gtactctggc

	50161	tgcagactga	cctctctcag	gtctgagaag	gatggagcag	ccactggagt	ggatgccatc
	50221	tgcattccatc	atcttgaccc	caaaagccct	ggactcaaca	gagagcggct	gtactgggag
	50281	ctgagccaac	tgaccaatgg	catcaaagag	ctgggcccct	acaccctgga	caggaacagt
	50341	ctctatgtca	atggtttcac	ccatcggacc	tctgtgcccc	ccaccagcac	tccctgggacc
5	50401	tccacagtgg	accttggaa	ctcagggact	ccattctccc	tcccaagccc	cgcaactgct
	50461	ggccctctcc	tggtgctggt	caccctcaac	ttcaccatca	ccaacctgaa	gstatgaggag
	50521	gacatgcatc	gccttggttc	caggaagtgc	aacaccactg	agagggtcct	gcagactctg
	50581	cttgggtccta	tggtcaagaa	caccagtgtt	ggccttctgt	actctggctg	cagactgacc
	50641	ttgctcaggt	ccgagaagga	tggagcagcc	actggagtgg	atgccatctg	caccaccgt
10	50701	cttgacccca	aaagccctgg	actgnacagn	gagcngctnt	actgggagct	nagccanctg
	50761	accaannnca	tcnngagct	gggnccctac	accctggaca	ggnacagtct	ctatgtcaat
	50821	ggttttcacc	atcnganctc	tgngcccacc	accagcactc	ctgggacctc	cacagtgnac
	50881	ntnggnacct	cngggactcc	atcctccntc	ccngccnca	cancnntcgn	ccctctcctg
	50941	ntnccnttca	ccntcaactt	naccatcacc	aacctgcant	angnganna	catgcnnnc
15	51001	ccnggntcca	ggaagttaa	caccacngag	agagtcttcc	agggctctgt	caggcctgtg
	51061	ttcaagaaca	ccagtgttgg	ccctctgtac	tctggctgca	gactgacctt	gctcaggccc
	51121	aagaaggatg	gggcagccac	caaagtggat	gccatctgca	cctaccgccc	tgatcccaaa
	51181	agccctggac	tggacagaga	gcagctatac	tgggagctga	gccagctaac	ccacagcatc
	51241	actgagctgg	gcccctacac	ccaggacagg	gacagtctct	atgtcaatgg	cttccaccat
20	51301	cggagctctg	tgccaaccac	cagtattcct	gggacctctg	cagtgcacct	ggaaaccact
	51361	gggactccat	cctccttccc	cggccacaca	gagcctggcc	ctctcctgat	accattcact
	51421	ttcaacttta	ccatcaccaa	cctgcgttat	gaggaanaa	tgcaaacacc	tggttccagg
	51481	aagttcaaca	ccacggagag	ggttctgcag	ggtctgtctc	cgcccttggt	caagaacacc
	51541	agtgttgccc	ctctgtactc	tggtctgcag	ctgaccttgc	tcagacctga	gaagcaggag
25	51601	gcagccactg	gagtggacac	catctgtacc	caccgcgttg	atcccatcgg	acctggactg
	51661	gacagagagc	ggctatactg	ggagctgagc	cagctgacca	acagcatcac	agagctggga
	51721	ccctacaccc	tgatagggga	cagtctctat	gtcgatggct	tcaacccttg	gagctctgtg
	51781	ccaaccacca	gcactcctgg	gacctccaca	gtgcacctgg	caacctctgg	gactccatcc
	51841	cccctgcctg	gccacacagc	ccctgtccct	ctcttgatac	cattcaccct	caactttacc
30	51901	atcaccgacc	tgcatatga	agaaaacatg	caacaccctg	gttccaggaa	gttcaacacc
	51961	acggagaggg	ttctgcaggg	tctgtccaag	cccttgttca	agagcaaccg	cggtggccct
	52021	ctgtactctg	gctgcagact	gaccttgctc	agacctgaga	aacatggggc	agccactgga
	52081	gtggacgcca	tctgcacctc	ccgcttgat	cccactggtc	ctggactgga	cagagagcgg
	52141	ctatactggg	agctgagcca	gctgaccaac	agcatcacag	agctgggacc	ctacaccctg
35	52201	gatagggaca	gtctctatgt	caatggcttc	aaccttgga	gctctgtgcc	aaccaccagc
	52261	actcctggga	cctccacagt	gcacctggca	acctctggga	ctccatcctc	cctgcctggc
	52321	cacacaactg	ctggccctct	cctggtgccc	ttcacctca	acttcacat	caccaacctg
	52381	aagtacgagg	aggacatgca	ttgcctgggc	tcagggaagt	tcaaacaccac	agagagagtc
	52441	ctgcagagtc	tgcatgggtc	catgttcaag	aacaccagtg	ttggccctct	gtactctggc
40	52501	tgcagactga	ccttgctcag	gtccgagaag	gatggagcag	ccactggagt	ggatgccatc
	52561	tgcacccacc	gtcttgaccc	caaaagccct	ggactgnaca	gngagcngct	ntactgggag
	52621	ctnagccanc	tgaccaannn	catcnngag	ctgggnccct	acaccctgga	caggnacagt
	52681	ctctatgtca	atggtttcac	ccatcnganc	tctgngccca	ccaccagcac	tctgggacc
	52741	tccacagtgn	acntnggnac	ctcngggact	ccatcctccn	tccccngccn	cacancnnt
45	52801	gnccctctcc	tgntnccntt	caccntcaac	tnnaccatca	ccaacctgca	ntangnggan
	52861	nacatgcnnn	nccnnggntc	caggaagtgc	aacaccacng	agnngtntct	gcagggtctg
	52921	ctnnnncccn	tnttcaagaa	cnccagtgtg	ggcctctgtg	actctggctg	cagactgacc
	52981	tnnctcaggn	cngagaagna	tggngcagcc	actggantgg	atgccatctg	cancaccnnc
	53041	cntnanccca	aaagnccctg	actgnacagn	gagcngctnt	actgggagct	nagccanctg
50	53101	accaacagca	tcacagagct	gggaccttac	accctggata	gggacagtct	ctatgtcaat
	53161	ggttttcacc	atcgaaagctc	tatgcccacc	accagtattc	ctgggacctc	tgcagtgcac
	53221	ctggaaacct	ctgggactcc	agcctccctc	cctggccaca	cagccctggg	ccctctcctg
	53281	tgccatttca	ccctcaactt	cactatcacc	aacctgcagt	atgaggagga	catgctgcac
	53341	cctggttcca	ggaagttaa	caccacggag	agagtctctg	agggctctgt	caagcccttg
55	53401	ttcaagagca	ccagtgttgg	ccctctgtac	tctggctgca	gactgacctt	gctcaggcct
	53461	gaaaaacgtg	gggcagccac	cggcgtggac	accatctgca	ctcaccgctc	tgacctctca
	53521	aaccttgagc	tgncacngna	gcnctnttac	tgggagctna	gccanctgac	caannnctc
	53581	nnngagctgg	gnccctacac	cctggacagg	nacagtctct	atgtcaatgg	tttcaaccat
	53641	cnganctctg	ngcccaccac	cagcactcct	gggacctcca	cagtgnacnt	nggnacctcn
60	53701	gggactccat	cctcctccc	cngccncaca	nonnctgncc	ctctcctgnt	ncnttccacc
	53761	ntcaacttca	ccatcaccaa	cctgcantn	gnggannaca	tgcnnccccc	nggntccagg

	53821	aagttcaaca	ccacngagng	ngtnctgcag	ggtctgctnn	nncccntntt	caagaacncc
	53881	agtgtnggcc	ntctgtactc	tggtctgcaga	ctgacctnnc	tcaggncnga	gaagnatggn
	53941	gcagccactg	gantggatgc	catctgcanc	caccnnctn	anccaaaag	ncctggactg
	54001	nacagngagc	ngctntactg	ggagctnagc	canctgacca	annncatcnn	ngagctgggn
5	54061	ccctacaccc	tggacaggna	cagtctctat	gtcaatgggt	ttcaccctcg	gagctctgtg
	54121	ccaaccacca	gcactcctgg	gacctccaca	gtgcacctgg	caacctctgg	gactccatcc
	54181	tccctgcctg	gccacacagc	ccctgtccct	ctcttgatac	cattcaccct	caactttacc
	54241	atcaccaacc	tgcattatga	agaaaacatg	caacaccctg	gttcaggaa	gttcaacacc
	54301	acggagcggg	tcctgcaggg	tctgcttggt	cccatgttca	agaacacaag	tgtcggcctt
10	54361	ctgtactctg	gctgcagact	gacctgtctc	aggcctgaga	agaatggggc	agccactgga
	54421	atggatgcca	tctgcagcca	ccgtcttgac	cccaaaagcc	ctggactgna	cagngagcng
	54481	ctntactggg	agctnagcca	nctgaccaan	nncatcnng	agctgggncc	ctacaccctg
	54541	gacaggnaca	gtctctatgt	caatggtttc	accatcnaga	nctctgngcc	caccaccatg
	54601	actcctggga	ccctccacagt	gnacntnggn	acctcnggga	ctccatcctc	cntccccngc
15	54661	cncacancnn	ctgnccctct	cctgntnccn	ttcacctnca	acttnaccat	caccaacctg
	54721	cantangngg	annacatgcn	nncccnnggn	tccaggaagt	tcaacaccac	ngagngngtn
	54781	ctgcagggtc	tgctnnnncc	cntnttcaag	aacnccagtg	tnggcctct	gtactctggc
	54841	tgcagactga	cctnnctcag	gncngagaag	natggngcag	ccactggant	ggatgccatc
	54901	tgcancacc	nnctnanc	caaaagncct	ggactgnaca	gngagcngct	ctatggggag
20	54961	ctnagccanc	tgaccaannn	catcnngag	ctgggncct	acaccctgga	caggnacagt
	55021	ctctatgtca	atggtttcac	ccatcagaac	tctgtgccca	ccaccagtac	tcctggggacc
	55081	tccacagtgt	actgggcaac	cactgggact	ccatcctcct	tccccggcca	cacagagcct
	55141	ggccctctcc	tgataccatt	cactttcaac	tttaccatca	ccaacctgca	ttataggaa
	55201	aacatgcaac	accctggttc	caggaagttc	aacaccacgg	agagggttct	gcagggtctg
25	55261	ctcacgccct	tgttcaagaa	caccagtgtt	ggccctctgt	actctggctg	cagactgacc
	55321	ttgctcagac	ctgagaagca	ggaggcagcc	actggagtgg	acaccatctg	taccaccgc
	55381	gttgatccca	tcggacctgg	actgnacagn	gagcngctnt	actgggagct	nagccancctg
	55441	accaannnca	tcnnngagct	gggnccctac	accctggaca	ggnacagctc	ctatgtcaat
	55501	ggtttcacc	atcngancct	tgngcccacc	accagcactc	ctgggacctc	cacagtgnac
30	55561	ntnggnacct	cngggactcc	atcctcctc	cccngccnca	cancnnetgn	ccctctcctg
	55621	ntnccnttca	ccntcaactt	naccatcacc	aacctgcant	angngganna	catgcnncnc
	55681	ccnggntcca	ggaagttcaa	caccacngag	ngngtnctgc	agggtctgct	nnnnccntn
	55741	ttcaagaacn	ccagtgtngg	cctctgttac	tctggctgca	gactgacctn	nctcaggncc
35	55801	gagaagnatg	gngcagccac	tggantggat	gccatctgca	nccacnncn	tnancccaaa
	55861	agnctggac	tnacagnga	gcngctntac	tgggagctna	gccanctgac	caannncatc
	55921	nnngagctgg	gnccctacac	cctggacagg	nacagtctct	atgtcaatgg	tttcaaccat
	55981	cggagctctg	tgccaaccac	cagcagtcct	gggacctcca	cagtgcacct	ggcaacctct
	56041	gggactccat	cctccctgcc	tgccacacaca	gcccctgtcc	ctctcttgat	accattcacc
40	56101	ctcaacttta	ccatcaccaa	cctgcattat	gaagaaaaca	tgcaacaccc	tggttccagg
	56161	aagttcaaca	ccacggagag	ggttctgcag	ggtctgctca	agcccttggt	caagagcacc
	56221	agtgttgccc	ctctgtactc	tggtctgcaga	ctgaccttgc	tcagacctga	gaaacatggg
	56281	gcagccactg	gagtggacgc	catctgcacc	ctccgccttg	atcccaactg	tcctggactg
	56341	nacagngagc	ngctntactg	ggagctnagc	canctgacca	annncatcnn	ngagctgggn
45	56401	ccctacaccc	tggacaggna	cagtctctat	gtcaatgggt	tcacccatcn	ganctctgng
	56461	cccaccacca	gcactcctgg	gacctccaca	gtgnacntng	gnacctcngg	gactccatcc
	56521	tcntccccn	gccncacanc	nnctgnccct	ctcctgntnc	onttcaccnt	caacttnacc
	56581	atcaccaacc	tgcantangn	ggannacatg	cnnncccng	gntccaggaa	gttcaacacc
	56641	acngagngng	tnctgcaggg	tctgctnnnn	ccntnttca	agaacnccag	tgtnggcnt
50	56701	ctgtactctg	gctgcagact	gacctnnctc	aggncngaga	agnatggngc	agccactgga
	56761	ntggatgcca	tctgcancca	ccnnctnann	cccaaaagnc	ctggactgna	cagngagcng
	56821	ctntactggg	agctnagcca	nctgaccaan	nncatcnng	agctgggncc	ctacaccctg
	56881	gacaggnaca	gtctctatgt	caatggtttc	acccatcgga	cctctgtgcc	caccaccagc
	56941	actcctggga	cctccacagt	gcacctggga	acdtctggga	ctccatcctc	cctgacctggc
55	57001	cacacagccc	ctgtccctct	cttgatacca	ttcacctca	actttaccat	caccaacctg
	57061	cagtatgagg	aggacatgca	tcgccttgga	tctaggaagt	tcaacaccac	agagagggtc
	57121	ctgcagggtc	tgcttagtcc	cattttcaag	aactccagtg	ttggccctct	gtactctggc
	57181	tgcagactga	cctctctcag	gcccagagaag	gatggggcag	caactggaat	ggatgctgtc
	57241	tgccctctacc	accctaatac	caaaagacct	gggctggaca	gagagcagct	gtactgagag
	57301	ctaagccagc	tgaccacaaa	catcactgag	ctgggcccct	acagcctgga	cagggacagt
60	57361	ctctatgtca	atggtttcac	ccatcagaac	tctgtgccca	ccaccagtac	tcctggggacc
	57421	tccacagtgt	actgggcaac	cactgggact	ccatcctcct	tccccggcca	cacancnct

57481 gnccctctcc tgntnccntt caccntcaac ttnaccatca ccaacctgca ntangnggan
 57541 nacatgcnnn nccnnggntc caggaagttc aacaccacng agnngntnct gcagggtctg
 57601 ctnnnncccn tnttcaagaa cncagtgtn ggccntctgt actctggctg cagactgacc
 57661 tnnctcaggn cngagaagna tggngcagcc actggantgg atgccatctg canccacnn
 57721 cntnanccca aaagnccctgg actgnacagn gagcngctnt actgggagct nagccanctg
 57781 accaannnca tcnnngagct gggncctac accctggaca ggnacagtct ctatgtcaat
 57841 ggtttcaccc attggagctc tgggctcacc accagcactc ctgggacttc cacagttgac
 57901 cttggaacct cagggactcc atccccctg cccagcccca caactgctgg ccctctcctg
 57961 gtgccattca ccctaaactt caccatcacc aacctgcagt atgaggagga ctatgcctgc
 58021 cctggatcta ggaagttcaa cgccacagag agggctcctgc agggctctgct tagtcccata
 58081 ttcaagaaca ccagtgttg ccctctgtac tctggctgca gactgacctt gctcagacct
 58141 gagaagcagg aggcagccac tggagtggac accatctgta cccaccgctg tgatcccatc
 58201 ggacctggac tgnacagna gcnctntac tgggagctna gccanctgac caannncatc
 58261 nnnagctgg gncctacac cctggacagg nacagtctct atgtcaatgg ttccacctat
 58321 cnganctctg ngcccaccac cagcactcct gggacctcca cagtgnacnt nggnacctcn
 58381 gggactccat cctccntccc cngccnaca ncnmctgncc ctctcctgnt nccnttcacc
 58441 ntcaacttna ccatcaccaa cctgcantan gnggannaca tgcnnncncc nggntccagg
 58501 aagttcaaca ccaengagng ngtnctgcag ggtctgctnn nnccntntt caagaacncc
 58561 agtgtnggcc ntctgtact tggctgcaga ctgacctnnc tcaggncnga ctagcnctgc
 58621 gcagccactg gantggatgc catctgcanc caccnncntn ancccaaaag nccctggactg
 58681 nacagnagc ngctntactg ggagctnagc canctgacca annncatcnn ngagctgggn
 58741 ccctacaccc tggacaggna cagtctctat gtcaatgggt tcaacctcgc gagctttggg
 58801 ctacaccaca gactccttg gacttccaca gttgacctg gaacctcagg gactccatcc
 58861 cccgtcccca gccccacaac tgctggcct ctctgggtgc catccacct aaacttcacc
 58921 atcaccaacc tgcagtatga ggaggacatg catcgccctg gctccaggaa gttcaacacc
 58981 acggagaggg tcttccagg tctgtctacg ccttggttca ggaacaccag tgtcagctct
 59041 ctgtactctg gttgcagact gacctgtctc aggcctgaga aggatggggc agccaccaga
 59101 gtggatgctg tctgcacca tctgtctgac cccaaaagcc ctggactgna cagnagcng
 59161 ctntactggg agctnagcca nctgaccaan nncatcnng agctgggncc ctacacctg
 59221 gacaggnaca gtctctatgt caatggtttc acccatcnga nctctgngcc caccaccagc
 59281 actcctggga cctccacagt gnacntnggn acctonggga ctccatctc cntcccncc
 59341 cncacancnn ctgnccctct cctgntnccn ttcacntca acttnaccat caccacactg
 59401 cantangng annacatgcn ncncccnngn tccaggaagt tcaacaccac ngagnngntn
 59461 ctgcagggtc tgctnnnncc cntnttcaag aacnccagt tnggcnctct gtactctggc
 59521 tgagactga cctnnctcag gncngagaag natggngcag ccaactggant ggatgccatc
 59581 tgcanccacc ncnctnancc caaaagncc ggactgnaca gngagcngct ntactgggag
 59641 ctngaccanc tgaccaannn catcnngag ctgggnccct acacctggga caggnacagt
 59701 ctctatgtca atggtttcac ccattggatc cctgtgcca ccagcagac tccctgggacc
 59761 tccacagtgg accttgggtc agggactcca tccctccctc ccagcccccac aactgctggc
 59821 cctctcctgg taccattcac cctcaacttc accatcacca acctgcagta tggggaggac
 59881 atgggtcacc ctggctccag gaagtccaac accacagaga gggctcctgca gggctgctt
 59941 ggtcccatat tcaagaacac cagtgttggc cctctgtact ctgggtgcag actgacctct
 60001 ctccaggtccg agaaggatgg agcagccact ggagtggatg ccactgtcat ccactcatct
 60061 gacccccaaa gccctggact gnacagnag cngctntact gggagctnag ccacttgacc
 60121 aannncatcn nngagctggg nccctacacc ctggacaggn acagtctcta tgtcaatggg
 60181 ttcacccatc nganctctgn gcccaaccac agcactcctg ggacctccac agtgnacntn
 60241 ggnacctcng ggactccatc ctccntcccc ngccnccan cnnctgnccc tctcctgntn
 60301 ccttccaccn tcaacttnac catcaccaac ctgcantag nggannacat gcnnncnccn
 60361 ggntccagga agttcaacac cacngagnn gtntctcagg gtctgctnnn nccntnttc
 60421 aagaacncca gtgtnggcn tctgtactct ggctgcagac tgacctnntc caggncngag
 60481 aagnatggng cagccactgg antggatgcc atctgcanc accnncntna ncccaaaagn
 60541 cctggactgn acagnagcn gctntactgg gagctnagcc anctgaccaa nnnatcnnn
 60601 gagctgggnc cctacaccct ggacaggnc agtctctatg tcaatgggtt caccatcag
 60661 acctttgcgc ccaacaccag cactcctggg acctccacag tggacctggg gacctcagg
 60721 actccatcct cctcccccag cctacatct gctggccctc tctgggtgac attcaccctc
 60781 aacttcacca tcaccaacct gcagtacgag gaggacatgc atcaccagg ctccaggaag
 60841 ttcaacacca cggagcgggt cctgcagggt ctgcttgggt ccatgttcaa gaacaccagt
 60901 gtcggccttc tgtactctgg ctgcagactg acctgtctca ggcttgagaa gaatggggca
 60961 gccaccagag tggatgctgt ctgcacccat cgtcctgacc ccaaaagccc tggactgnac
 61021 agngagcngc tntactggga gctnagccan ctgaccaann ncatcnnga gctgggnccc
 61081 tacacctgg acaggnacag tctctatgtc aatggtttca ccatcnngan ctctgngccc

	61141	accaccagca	ctcctgggac	ctccacagt	nacntnggna	cctcngggac	tccatcctcc
	61201	ntccccngcc	ncacagcccc	tgccccctc	ttgataccat	tcacccctcaa	ctttaccatc
	61261	accaacctgc	attatgaaga	aaacatgcaa	caccctgggt	ccaggaagtt	caacaccacg
	61321	gagagggttc	tgcaagggtc	gctcaagccc	ttgttcaaga	gcaccagcgt	tggccctctg
5	61381	tactctggct	gcagactgac	cttgctcaga	cctgagaaac	atggggcagc	cactggagtg
	61441	gacgccatct	gcaccctccg	ccttgatccc	actggctcctg	gactggacag	agagcggcta
	61501	tactgggagc	tgagccagct	gaccaacagc	gttacagagc	tgggccccta	caccctggac
	61561	agggacagtc	tctatgtcaa	tggtttcacc	cagcggagct	ctgtgccaac	caccagtatt
	61621	cctgggacct	ctgcagtcca	cctggaaacc	tctgggactc	cagcctccct	ccctggccac
10	61681	acagccccctg	gccccctcct	ggtgccattc	accctcaact	tcactatcac	caacctgcag
	61741	tatgagggtg	acatgcgtca	ccctgggttc	aggaagtcca	acaccacgga	gagagtctctg
	61801	cagggctctgc	tcaagccctt	gttcaagagc	accagtgttg	gccctctgta	ctctggctgc
	61861	agactgacct	tgctcaggcc	tgaaaaacgt	ggggcagcca	ccggcggtga	caccatctgc
	61921	actcaccgcc	ttgacctctc	aaaccttgga	ctggacagag	agcagctata	ctgggagctg
15	61981	agcaaactga	cccgtggcat	catcgagctg	ggccccctacc	tcctggacag	aggcagctctc
	62041	tatgtcaatg	gtttcaccga	tcggaaacttt	gtgcccatac	ccagcactcc	tgggacctcc
	62101	acagtacacc	taggaacctc	tgaaaactcca	tcctccctac	ctagacccat	agtgcctggc
	62161	cctctcctgg	tgccattcac	cctcaacttc	accatcacca	acttgcagta	tgaggaggcc
	62221	atgcgacacc	ctggctccag	gaagttcaat	accacggaga	gggtctctaca	gggtctgctc
20	62281	aggcccttgt	tcaagaatac	cagtatcggc	cctctgtact	ccagctgcag	actgaccttg
	62341	ctcaggccag	agaaggacaa	ggcagccacc	agagtggatg	ccatctgtac	ccaccacctt
	62401	gacctcaca	gccccggact	gaacagagag	cagctgtact	gggagctgag	ccagctgacc
	62461	cacggcatca	ctgagctggg	cccctacacc	ctggacaggg	acagtctcta	tgtcgatggg
	62521	ttcactcatt	ggagcccatc	agcaccacc	agcactcctg	ggacctccat	agtgaacctg
25	62581	ggaacctctg	ggatcccacc	ttccctccct	gaaactacan	cnnctgnccc	tctcctgntn
	62641	ccnttcaccn	tcaacttnac	catcaccaac	ctgcantang	nggannacat	gcnnccnccn
	62701	ggntccagga	agttcaacac	cacngagagg	gttctgcagg	gtctgctcaa	gcccttggtc
	62761	aagagcacca	gtgttgcccc	tctgtattct	ggctgcagac	tgacctgtgt	caggccctgag
	62821	aaggacggag	tagccaccag	agtggaagcc	atctgcaccc	accgacctga	ccccaaaatc
30	62881	cctgggctag	acagacagca	gctatactgg	gagctgagcc	agctgaccca	cagcatcact
	62941	gagctgggac	cctacaccct	ggatagggac	agtctctatg	tcaatgggtt	caccacagcg
	63001	agctctgtgc	ccaccaccag	cactcctggg	actttcacag	tacagccgga	aacctctgag
	63061	actccatcat	ccctccctgg	ccccacagcc	actggccctg	tcctgctgcc	attcacccctc
	63121	aattttacca	tcactaacct	gcagctatgag	gaggacatgc	atcgccctgg	ctccagggaag
35	63181	ttcaacacca	cggagagggt	ccttcagggt	ctgcttatgc	ccttggttcaa	gaacaccagt
	63241	gtcagctctc	tgtactctgg	ttgcagactg	accttgctca	ggcctgagaa	ggatggggca
	63301	gccaccagag	tggatgctgt	ctgcacccat	cgtcctgacc	ccaaaagccc	tggactggac
	63361	agagagcggc	tgtactggaa	gctgagccag	ctgaccacag	gcatactgta	gctgggcccc
	63421	tacaccctgg	acaggcacag	tctctatgtc	aatgggttca	cccatcagag	ctctatgacg
40	63481	accaccagaa	ctcctgatac	ctccacaatg	cacctggcaa	cctcgagaac	tcacagcctcc
	63541	ctgtctggac	ctacgaccgc	cagccctctc	ctgggtgctat	tcacaattaa	cttcaccatc
	63601	actaacctgc	ggtatgagga	gaacatgcat	caccctggct	ctagaaagtt	taacaccacg
	63661	gagagagtcc	ttcagggtct	gctcaggcct	gtgttcaaga	acaccagtgt	tggccctctg
	63721	tactctggct	gcagactgac	cttgctcagg	cccaagaagg	atggggcagc	caccaaagtg
45	63781	gatgccatct	gcacctaccg	ccctgatccc	aaaagccctg	gactggacag	agagcagcta
	63841	tactgggagc	tgagccagct	aaccacagc	atcactgagc	tgggccccta	caccctggac
	63901	agggacagtc	tctatgtcaa	tggtttcaca	cagcggagct	ctgtgcccac	cactagcatt
	63961	cctgggaccc	ccacagtggg	cctgggaaca	tctgggactc	cagttttcta	acctgggtccc
	64021	tgggctgcca	gccccctcct	ggtgctattc	actctcaact	tcaccatcac	caacctgcgg
50	64081	tatgaggaga	acatgcagca	ccctggctcc	aggaagtcca	acaccacgga	gagggctcctt
	64141	cagggcctgc	tcagggtccc	gttcaagagc	accagtgttg	gccctctgta	ctctggctgc
	64201	agactgactt	tgctcaggcc	tgaaggagat	gggacagcca	ctggagtgga	tgccatctgc
	64261	accaccacc	ctgaccccaa	aagccctagg	ctggacagag	agcagctgta	ttgggagctg
	64321	agccagctga	cccacaatat	cactggagctg	ggccactatg	ccctggacaa	cgacagcctc
55	64381	tttgtcaatg	gtttcactca	tcggagctct	gtgtccacca	ccagcactcc	tgggaccccc
	64441	acagtgtatc	tgggagcatc	taagactcca	gcctcgatat	ttggcccttc	agctgccagc
	64501	catctcctga	tactattcac	cctcaacttc	accatcacta	acctgcggta	tgaggagaa
	64561	atgtggcctg	gctccaggaa	gttcaacact	acagagaggg	tccttcaggg	cctgctaagg
	64621	cccttggtca	agaacaccag	gtttggccct	ctgtactctg	gctccaggct	gccttgctc
60	64681	aggccagaga	aagatggggg	agccaccgga	gtggatgcca	tctgcaccca	ccgcccctgac
	64741	cccacaggcc	ctgggctgga	cagagagcag	ctgtatttgg	agctgagcca	gctgacccac

5 64801 agcatcactg agctgggccc ctacacactg gacagggaca gtctctatgt caatgggttc
 64861 acccatcgga gctctgtacc caccaccagc accgggggtg tcagcgagga gccattcaca
 64921 ctgaacttca ccatcaacaa cctgcgctac atggcggaca tgggccaacc cggtccctc
 64981 aagttcaaca tcacagacaa cgtoatgaag cacctgctca gtcccttgtt ccagaggagc
 65041 agcctgggtg cacggtacac aggctgcagg gtcacgcac taaggctctgt gaagaacggt
 65101 gctgagacac ggggtggacct cctctgcacc tacctgcagc ccctcagcgg ccaggtctg
 65161 cctatcaagc aggtgttcca tgagctgagc cagcagaccc atggcatcac ccggctgggc
 65221 ccctactctc tggacaaaga cagcctctac cttaacggtt acaatgaacc tggcttagat
 10 65281 gagcctccta caactcccaa gccagccacc acattcctgc ctctctgtc agaagccaca
 65341 acagccatgg ggtaccacct gaagacctc acactcaact tcaccatctc caatctccag
 65401 tattcaccag atatgggcaa gggctcagct acattcaact ccaccgaggg ggtccttcag
 65461 cacctgctca gaccttgtt ccagaagagc agcatgggccc ccttctactt ggggttgccaa
 65521 ctgatctccc tcaggcctga gaaggatggg gcagccactg gtgtggacac cacctgcacc
 65581 taccacctg accctgtggg cccgggctg gacatacagc agctttactg ggagctgagt
 15 65641 cagctgaccc atggtgtcac ccaactgggc ttctatgtcc tggacagggg tagcctcttc
 65701 atcaatggct atgcacccca gaatttatca atccggggcg agtaccagat aaatttccac
 65761 attgtcaact ggaacctcag taatccagac cccacatcct cagagtagat caccctgctg
 65821 agggacatcc aggacaaggc caccacactc tacaaaggca gtcaactaca tgacacattc
 65881 cgcttctgcc tggtcaccaa cttgacgatg gactccgtgt tggctcactgt caaggctattg
 20 65941 ttctcctcca atttggaccc cagcctgggtg gagcaagtct ttctagataa gacctgaat
 66001 gcctcattcc attggctggg ctccacctac cagttgggtg acatccatgt gacagaaatg
 66061 gagtcatcag tttatcaacc aacaagcagc tccagcacc agcacttcta cctgaatttc
 66121 accatcacca acctaccata ttcccaggc aaagcccagc caggcaccac caattaccag
 66181 aggaacaaaa ggaatattga ggatgcgtc aaccaactct tccgaaacag cagcatcaag
 25 66241 agttattttt ctgactgtca agtttcaaca ttcaggtctg tcccaacag gcaccacacc
 66301 ggggtggact cctgtgttaa cttctcgcca ctggctcgga gtagtagacag agttgccatc
 66361 tatgaggaat ttctgcggtt gaccgggaat ggtacccagc tgcagaactt caccctggac
 66421 aggagcagtg tccttgtgga tgggtattct ccaacagaa atgagccctt aactgggaat
 66481 tctgaccttc cttctgggc tgtcatctc atcggcttgg caggactcct gggactcatc
 30 66541 acatgcctga tctgcggtgt cctggtgacc acccgccggc ggaagaagga aggagaatac
 66601 aacgtccagc aacagtgcac aggtactac cagtacacac tagacctgga ggatctgcaa
 66661 tgactggaaac ttgccgggtg ctgggggtgc tttccccag ccagggtcca aagaagcttg
 66721 gctggggcag aaataaacca tattggtcgg aaaaaaaaaa aaaaa

35 SEQ ID NO. 3

hk5 amino acid

40 MATARPPWMWVLCALITALLGVTEHVLANNVSDHPSNTVPSSNQDLGAGAGEDARSDDSSSRIINGS
 CDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVRLGHYSLSFVYESGQQMFQGVKSIPHPG
 YSHPGHSNDLMLIKLNRIRPTKDVRIINVSSHCPGAGTKCLVSGWGTTKSPQVHFPPKVLQCLNISVLSQKR
 CEDAYPRQIDDTMFCAGDKAGRSDSCQGDGSGGPVVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQET
 IQANS

45 SEQ ID NO. 4

KLK5 CDS

50 ggtgtctgtg cgtcctgcac ccacatcttt ctctgtcccc tccttgccct gtctggaggc
 tgctagactc ctatcttctg aattctatag tgccctgggtc tcagcgcagt gccgatgggtg
 gcccgctcctt gtggttcctc tctacttggg gaaatcaggt gcagcggcca tggctacagc
 aagaccccc tggatgtggg tgctctgtgc totgatoaca gccttgcttc tgggggtcac
 agagcatggt ctgcacaaca atgatgtttc ctgtgacccac ccctctaaca ccgtgccttc
 tgggagcaac caggacctgg gagctggggc cggggaagac gcccggtcgg atgacagcag
 cagccgcac atcaatggat ccgactgcga tatgcacacc cagccgtggc aggccgcgt
 gttgctaagg cccaaccagc tctactgcgg ggccggtgttg gtgcatccac agtggctgct
 55 cacggccgcc cactgcagga agaaagtgtt cagagtccgt ctcgccact actccctgtc
 accagtttat gaatctgggc agcagatggt ccagggggtc aaatccatcc cccaccctgg
 ctactccac cctggccact ctaacgacct catgtcctc aaactgaaca gaagaattcg
 tcccactaaa gatgtcagac ccatcaacgt ctctctcat tgtccctctg ctgggacaaa


```

      gtgcttgggtg tctggctggg ggacaaccaa gagcccccaa gtgcacttcc ctaaggtcct
      ccagtgccttg aatatcagcg tgctaagtca gaaaagggtgc gaggatgctt acccgagaca
      gatagatgac accatgttct ggcgcgggtga caaagcaggt agagactcct gccaggggtga
      5  ttctgggggg cctgtggtct gcaatggctc cctgcaggga ctctgcaagt tcaccaagtg
      cccttgtgcc cggcccaaca gaccgggtgt ctacacgaac ctctgcaagt tcaccaagtg
      gatccaggaa accatccagg ccaactcctg agtcatccca ggactcagca caccggcatc
      cccacctgct gcagggacag ccctgacact cctttcagac cctcattcct tcccagagat
      gttgagaatg ttcattctct cagcccctga ccccatgtct cctggactca ggggtctgctt
      ccccccacatt gggctgaccg tgtctctcta gttgaaccct gggaacaatt tccaaaactg
      10  tccagggcgg gggttgogtc tcaatctccc tggggcactt tcatcctcaa gctcagggcc
      catcccttct ctgcagctct gacccaaatt tagtcccaga aataaactga gaagtggaaa
      aaaaaaa

```

15 SEQ ID NO. 5

KLK5 nucleic acid

```

      gggcccagag tgaaggcaag agaaggagtt gagagctccc tctgcaaagt ggcttgagtc
      20  tcccctgcct aaaatgcagg gagagggagg cagaaagaca ggggaagagga aggggtgggg
      aagaaaagaga gagagagaga gagacagaat aacacaacta cagaaacaca gagagaacac
      acagagagcc tgggacacag ggacacacag agtcagagag aaaagagaag atagagaaag
      acacaaatgg agacacagag gtgtaaagaa agagagatta acagagtccc agatacacgc
      aaaggggcag aagcacagtt ttcaggggtgg tgtctatgat catcttcttt tttttttttt
      25  tttttttttt tttttgagac ggagtctcgc tctgtcgcgc aggctggagt gcagtggcgg
      gatctcgggt cactgcaagc tccgcctccc gggttcacgc cattctctg cctcagcctc
      ccaagtagct gggactacag gcgcccgcga ctacgcccgg ctaatttttt tgtattttta
      gtagagacgg ggtttcaccg ttttagccgg gatggcctcg atctcctgac ctctgtatcc
      gccgcctcgg gcctcccaaa gtgctgggat tacaggcgtg agccaccgcg cccggccatg
      atcatcttct tgaotatgct gatgtgacaa gtacctaaag ccatcagact ctacccttta
      30  aatatgcagt ttggggccagg caccgtggct catgcctgta attccagcac tttgggaggc
      agaggtgggt gaatcacttg aggcaggag tttgagacca gcctggccaa catggtgaaa
      ctctgtcttt actaaaaaaaa aaaaaaaaaa aaaaaaaatc agccgggtgt cgtggggcac
      acctgtaatc ccagctatgc tggaggctga ggcacgagag tcaattgaac cctggaggcg
      gaggttgagc tggggccgaga tcacatcacc gccctccagc ctggggcgaca gagcaagact
      35  ctgtctcaaa taataaata aacaaacgaa caagcagttt gttgtacctt agttatatct
      aaaaaaaaaa tgctgtcaac aaatagagca gaagtgaat aaaggaaaaa aaatgggcca
      agaactctaa ggtataattg acaaatcatt cagaaccttt aaaaaagaaa gaatcacaga
      ggcatagaaa gacagggagg aacagggaga cagaaacacc tgtggcccaa ggagaacaaa
      40  acaaggctoc taagacagac aggaggagag agagagagag tgagtgaag acagacagag
      aaaaagacag agagagagag acagagacag agagacagag agggcgagagg gatagaaaga
      gagagagggg tggagagaga cacgagatat tgagagagac tcagaaagat agccgaggga
      gaaccacaga gagatggaag aagactctga gaaaaaacca gagacaaaaga tggaaagagg
      agtatcgagg gtgaacagac agtggtggaa tgagcaaaat gcagagaaga aagcaagcaa
      45  tccaggcgcc aagaatagtg acccagagtt ggtgagaagc cagatcctta aggtcgggg
      aggcagggaa ggggctggcc tggcttccgg agaccctcc ccatctctcc ggccagggag
      gtagggagtg acattccgga ctgggtgggg ggtgctctgg ggggtggagt agggggagca
      ggaggagcta ttgctaaggc ccgataggca cctcattgcc cgggaatgtg cccagggag
      cagtgggtgg ttataactca ggcccgtgc ccagagccca ggaggaggca gtggccagga
      50  aggcacaggg ctgagaagtc tgcggtgag ctgggagcaa atccccacc cctacctgg
      gggacagggc aagtgaagc tggtaggggt ggctcagcag gcagggaagg agaggtgtct
      gtgcgtcctg caccacatc tttctctgtc cctccttgc cctgtctgga ggctgctaga
      ctcttatctt ctgaattcta tagtcctgg gtctcagcgc agtgccgatg gtggcccgtc
      cttgtgggtt ctctctacct ggggaaataa ggtaggggag ggaggggaag tgggttaagg
      gctccccgga tcgcctgggc ctcccaaccc tctgacatc cccatccagg tgcagcggcc
      55  atggctacag caagacccc ctggatgtgg gtgctctgtg ctctgatcac agccttgctt
      ctgggggtca caggtaacca gaactctggg tggggagggt tgtgggattg ggaggactgt
      ctctgcggca ctagagcgcc tgtccctgg ggaactgtgt gagcctgggc atgactccgg
      gaccgggtga atgtgagtct ctgtctgtac ttgtggttgt gcgatcgtat gtggccctgt

```

	gactgccacg	gtgtgtgtcg	gggaggggga	tgccttttcc	catatcaggt	gactgtgogg
	caggtggcac	tgaccctttg	aggctgtgtg	tgtggttttg	tgatttgttg	tgcatttaag
	atttgtgtgtg	gctccacagc	tgtgtgggtg	aatgcattga	gcactggggg	tgttcactgt
5	gtgtttggct	gtgtgtgggt	acttggcatt	gtatatgact	gcaggtatct	gcagttcctg
	tccctgaggt	cccgggattg	cgtgcaacaa	aagtggcat	caccatggaa	agctgtgact
	gtgtgctgct	tgcaggcgat	tatgtgattg	tggctgagtg	tgacgttatg	gatgcccgta
	tttgtgaccg	tgtgactacc	tgaagctctg	tgtagggttg	actgtatgtg	actgtgtgtg
	tctgtgtgag	gccgtgtaaa	tgctactgta	tgtgtgatgg	tgcagctgtg	tgtctggagt
	ttctgtctct	gcctggaggg	atagagggtg	caggggtagc	tatctctggg	agatgggtgc
10	caggtgactg	acttgcagtg	tgtgcctgtg	tgcagaagag	tatgtggcag	tctgaacatc
	tgtgcacaca	cggcatctgt	gcgtggcact	gagacactgt	ggatgagggt	gtgcgactcc
	gctaggctgc	ccgggagcgt	gtgtacctgg	agacagagct	gtatgttagc	tgcacctgtg
	gaggcaacat	gggogtgtct	gcagaactgc	gtgcgtgctt	ggctgttact	gctgtttgtg
	gcgtgggtct	tgggggtgagt	tcgtgaatga	tgggtgtgcc	agggccatca	gcaagggtaa
15	gaaccaggcc	gggogcggtg	gctcacgcct	gtaatcccag	ccctttggga	ggccgaggca
	ggcggtacac	ctgaggtcgg	gagatcgagg	ccagcctgac	caacatggag	aaccccgctc
	ctactaaaaa	tacaaaaaat	tagctgggtg	ggtggcgctg	gcctgtaatc	ccagctactc
	gggagactgg	ggcagaaaaa	tcgcttgaac	ccgggaggtg	gaggttgctg	tgagccgaga
	tcgogccatt	gcactccagc	ctgggcaaca	agagcgaaac	tcogtctcga	aagaaaaaaa
20	gaaaaaaaaa	agggtaaaga	ccagtgaatg	ggcacgggag	gactgatgat	ggagtggggc
	atgcattgtag	tctgttaggtc	tgtgtgtgag	aggaggagat	tgacaggatt	gagaaggcat
	gttttcattct	gagaatttcag	aaacctaggc	ctgctcttcc	cctccatgtg	gccccctaag
	ctgagccctt	ctttcctggt	cctgctttcg	gaacctagc	tcogcccatg	agctctgacc
	ccacctcctt	tcctcaacca	cgccccctag	ccagactcta	gtggaccccc	cctaaggcca
25	cacccctttg	ggccaggctc	cacccccctat	tctgtgggta	ccttctagaa	cccccttcaa
	agtcagagct	tttttttttt	tttttttgga	gacagtcttg	ctctctctcc	caggctggag
	tgcagtggcg	tgatctcggc	tcactgcaac	ctctgcctcc	caggttcaag	tgattctcgt
	gcctccacct	cctgagtagc	tgggattaca	ggtgcgcgcc	accacgcctg	gctaattttt
	gtgtcttttag	tagagacagg	gtttcacctt	gttgccaggg	ctgggtctca	actcccaacc
30	tcagggtgatc	cgcccacctc	ggcctcccag	agtgtctggg	ttacaggcgt	gagccaccgc
	ccccagccca	aagtcagagc	tctttatagg	agactctaac	atgtaaccct	gaccctggcc
	ctaactaagt	caattccaaa	ccccttcctg	cctccagccc	tgacccctct	cactgaggcc
	tgacccctct	tcttgagacc	agttccatcc	ctaaagccct	ggtctccctc	ccatccccag
	gctccagccc	ccacagcttt	ggcactaccc	ctgagcttgt	ccaggaatcc	tgtacccaat
35	ttttaccctca	catgtagtct	tagccaattc	caggaatctg	tgaggctccag	ttagagtcca
	gtaaccctac	ctgagcctgg	gctctgtcct	tgagcttgag	cctgggcttg	agaggtgcca
	ctcttattct	ccaggccctg	cccctgcccc	ctcagcatgt	cagacaccca	ccctctagct
	ggtctggcct	cttgagctctg	aaaccacccc	ccagcccaag	ccccgcctct	gagccccgcc
	caaccctatt	tcggttccca	gagcatgttc	tcgccaacaa	tgatgtttcc	tgtgaccacc
40	cctctaacac	cgtgcctctc	gggagcaacc	aggacctggg	agctggggcc	ggggaagacg
	cccgttcgga	tgacagcagc	agccgcacac	tcaatggatc	cgactgcgat	atgcacaccc
	agccgtggca	ggccgcgctg	ttgctaaggc	ccaaccagct	ctactgcggg	cggtgttg
	tgcattccaca	gtggctgctc	acggccgccc	actgcaggaa	gaagtgagtg	ggagtcccaa
45	gaggaggggt	ggtggggacg	gggaagtggg	ggtgggggtg	gggaagtggg	ggtgggggtg
	tcattggagg	gagggctggt	ggggacgggg	aagtgggggt	gggggtgtca	tggaagggtga
	gggttggttg	ggatgggttg	gggatgtggg	agcaggagga	ggtcgagttg	gggataggac
	taaggatgga	gttttgcggg	ggagcaagg	gggaggatga	ggttgagag	gggagagtgt
	tgtggtaggg	aatgggaagg	agccaaggat	gggttggttg	tggggttagg	agcatatatt
	tggtgaatgg	tttgggatgg	aggtggaatt	gggattggct	ttagaatttg	gggtgggtga
50	aaatcgggct	ggggtggaaa	tgaagatagc	atggagatag	ggttgagatt	gggagcagat
	atagaatgaa	ggatggggat	tggagttttg	ggtgggggtg	gagatgggtg	gatttgggct
	tgagaatgca	tatggtgatg	gcttctgggt	agggaaagaa	ttaggggttg	gaatgggatg
	gttttggaat	tgtgactggg	atggggacag	gcatgggatt	ggagaccaag	agggagttga
	ggatgggttg	gggaccgggg	gtggggatgg	gggtggggct	ggggctgggt	gtgggggttg
55	gattggcggt	ggacgtggag	atagagatca	gggttggttg	tgacctgcc	catcttctct
	agagttttca	gagtcctgtc	ogggcactac	tcctgttcac	cagtttatga	atctgggcag
	cagatgttcc	agggggtcaa	atccatcccc	caccctggct	actcccaccc	tgggcactct
	aacgacctca	tgctcatcaa	actgaacaga	agaattcgtc	ccactaaaga	tgtcagaccc
	atcaacgtct	cctctcattg	tcctctctgt	gggacaaagt	gcttggtgtc	tggctggggg

	acaaccaaga	gcccccaagg	tgagtgtcca	ggttcttctt	gataccgacc	catctctgcc
	gccttccatc	tttctccact	tctcattgtg	ttcctgtttg	acagtgcact	tccctaaggt
	cctccagtgc	ttgaatatca	gcgtgctaag	tcagaaaagg	tgcgaggatg	cttacccgag
5	acagatagat	gacaccatgt	tctgcgccgg	tgacaaagca	ggtagagact	cctgccagggt
	gaggacacct	ctctttattc	agcagataca	cactgagtgc	caactcggta	acatggagcg
	ttgccaaatt	ctgagaatcc	agcaattgcc	aagacagtca	ggacccctgt	tctcacagag
	ctcataccct	agagtagtgg	tgttttagtag	aaataatgct	gagctgctta	tgtcatttcc
	agtttttttag	tagccacatt	aaaacaggta	aaaaaggctg	ggcgagtggt	ctcacacctg
	taatcccagc	acttttgggag	gctgaggcag	gcagatcacc	tttgggtcagg	agtttgagac
10	tagcctggcc	aacatggcga	aactctgtct	ctaaaaaaa	atacaaaaa	tagcctggca
	tggtggcggg	cgctgtaat	ctcagctgct	caggaggccg	agacacaaga	atcacttaaa
	cccaggaggt	ggagggttgca	gtgagctgag	atcgtgccac	tcactccaac	ctgggagaca
	gagtgacact	tttgtctcaa	aaagaaaaaa	aaaaacaagt	aaaaaagaaa	cagggtgaagt
	taactttaat	aaccacaatgt	atcccaata	caatcatttc	aaagttaaat	taatataaaa
15	caattatgaa	tgagatactt	tacattcttt	tcttgttttc	atattaagtc	tttgaaagtg
	agtatatatg	ttatgctgac	agcacatctc	aatttggact	agctacattt	cagggtgtca
	gtggcctagca	gttactgtat	gttactgtat	tggatggcac	ggatctagag	ggaaagatca
	gggctgtttt	gtatggtttg	gcagggttg	cactgcataa	agataccata	tctaataagg
	gcactccgtg	ttacagatgt	cagttttggc	agttttcagg	cggtgtgtag	ttagtgtctt
20	tgtttcaaca	aaatctgtaa	tatgacagtt	ttctagcaag	tgctgtgtaa	atatcttgag
	gaaggaaaag	agaaatctgg	taggtatttt	tacaagagaa	tatttaatac	aggggattaa
	ttgcaaagct	gctggaaggg	ctggaggaac	aaagttaaaa	aataaaaaac	tctgtggtca
	agaatctgca	taaatagggc	aatttcagag	agtggttaaag	gttaacccca	aaataaaaaca
	tggttttagg	atagtaaaaca	ataagggccca	atattcaaaa	aggtggtcag	gggagcctcc
25	ttggagaggt	ggcatttgag	cagagaatgg	atgacacaaa	gaagctaaac	tcgtgaagtt
	taaggggaaa	gaaaaggcac	gtgcaaaggc	cctgaggcag	taaggaaatt	ggctgattca
	aagaagaaga	ggaaaccaat	gcaactggag	aacaaaagt	ggggcaacag	tagaaagtga
	cgctggaggt	gtaggcaggg	gcgaatgctc	tgcaagtatt	tcttgggtc	caacacagag
	cttccctatg	ttctaattgga	agctgtatct	gttgagggaag	acagaattta	aaatcaaaact
30	gttacatcaa	ccagcacccct	tctctgtatt	caggctccca	agggatctag	aaggacgtaa
	gttaacaagc	tctcattagc	aggggtgtgt	tttcaacagt	agttaggaag	ctggggattc
	aggagtactc	cagtcccatg	gctatgaaaa	gctcccccca	aattgtacaa	acctgacaaa
	tgcaacacct	ccccagctct	ccccatttct	tctctgtgcc	ctgggtgtgg	gggggtgggt
	tgcgaggggg	aaaactttta	acagaagaaa	gcacatctcg	gccgggcgtg	gtgggtcaca
35	cctgtaatcc	caacactttg	ggaggccgag	gcgggtggat	cactagggtc	ggagctggag
	accatcctgg	ctgacacggg	gaaacccctgt	ctctactaaa	aacacaaaaa	attagccggg
	cggtgtggca	ggcgccctgt	gtcccagcta	ctcgaggagg	tgaggcagga	gaatggcctg
	aaccggggag	gcggaacttg	cagtgaagcc	aggttgacc	actgcactcc	agcctgggca
	acacagttag	actccgtctc	aaaaaaaaaa	aaagaaaaga	aaagaaatca	catctcattc
40	aagtgtgtgg	attttaaact	atttagcctt	tctgtaggca	aggttagtat	cttgtttttc
	cagactcaaa	ggtgtttttt	gtttgttttt	ttcataccgg	tgtgtgtgtc	gggtgtggcc
	actaaaagct	acaagcaaga	aataataaca	actacaacaa	tactaatacc	aatagtataa
	aaataatagc	atctggctaa	ttgtctggca	ctgtttttaag	tggttttgc	gcctcagctc
	attaactcat	ttacctgtta	ttattggccc	tatttttaca	acaaggagcc	aaggctcaga
45	gcagttaaact	aacagcctct	caaaagaaac	tctgcagaga	tattaaattt	aaaaaataat
	gagagaaatt	aaaccacaag	aaagtgtgaa	tttagaggta	caggcagcta	agcttgtttg
	ctttgaaaca	gtgtctgcta	ctgggaaaaa	ggcaagtctt	ggctttccta	ataattgata
	ccaggactct	gtaattcata	ttttgcatgc	atgtaagtaa	gaaatgaagc	cgggtgcaat
	ggcacatgcc	agtaatccca	gcactctggg	agactgaagt	gggaagatca	cttgagctca
50	ggagttcaag	accagcctgg	gcaactaaaa	attaaaaaaa	taaaaatact	aattgttttt
	attttagtag	attttatcca	taccacttac	atcattattg	tagtatgtac	atatttattt
	cttttctttt	cttttctttt	cttttttgag	acggagtctc	gctctgtcac	ccaggctgga
	gtgcaatggc	accatatoag	ctcactgcag	catgcgcctc	ctgggttcaa	gcatttcttc
	cacctcagcc	tcccaagtag	ctgggataac	aggcaccac	caccatgcct	ggctattttt
55	ttttttccgt	agagatgggg	ttccaccatg	ttggccaggc	tggtcttgaa	ctcctgacct
	ccagtgtatct	gcctgcctog	gcctcccaaa	ttgctgggat	tacaggtgtg	agccacogtg
	cccagggtgg	agatagacat	ttctctctac	ctcaaacaga	ggtccactca	agctactttt
	cattttcttc	ataaatatta	gccgagtggc	tattttgcac	cagggaatgg	tccaggtgct
	gtggatatgg	catcaggcaa	aacagaccaa	aaacttctct	ccgcgtggac	ctcatgttcc

ccaagtggaa gacaggcaat aaagagatag ataaatatgt agtaaattaa aaaaaaaaaa
 aattagccgg gtgtggtggc ttgcacctgt agttccagct acttgggagg ctgaggtggg
 agaattgctt gagcccaaac gtttgaggct gcggaagcc atgactgcac tgctgcactc
 5 cagacagcag cctgggtgac aaagcaagac gtttttgtca gaaagaaaa aaaaagagac
 gaaggaggga aggagagaga aagggaaggaa ggaaggagaa agaaaggaa agaaagaaa
 gaaaggaagg aaggaaggag aaagaaagga agaaagagaa agaaagaaa agaaagaaa
 aaagaagaaa gaaaagagag aggaaggaag gaaagaagga aaagaggga aaaaatgact
 gttgaagagc agtgagtatt attataggag ggtaattata gggaggtatg ggggaattgaa
 10 gacaggaaac acaaattagt ccaagcgaat ggatttctat tgggagtgat tctgccccta
 gaagacactg gcaataccag gagacathtt tggttgtcac aactatatgg aggggcatta
 ctggcaacta atggatagat gccaaagtgt ctgttcaaca tgctatgatg cacacggcag
 gcctccacaa caaaccatta tccagcttca gatgccaca gtgccagat cgaggaaccc
 tcatccaggg gctgagaacc gtatttttgc agaaggagg tataaggatg ggttgggtgga
 15 gaatggggaa ggaaggtgtg tgtccagtaa gagaaataag gcctgcacag gctggagggg
 agagtggag agaaaggag gcggagagat acacgatgag ggagacaggc tggaaacagaa
 agtagagacg aagattcgag atgtggagag gaagggtcac agaccccccc gaaatgatgt
 gtggacaaca ggaatctgga agaggaagat ggagtggaga gtgacaaatg gggctctaaag
 gttgaacttg gaggccaggc atggtggctc acgcctgtaa tcccaacact ttggaggctg
 20 aggtggggcga atcacttgag gccaggagtt cgagaccagc ctggccaaca tggtgaaacc
 ccgtctctac aaaaaaata caaaaaatta gccgggtgtg gtgatggaca cctgtagtca
 cagctacttg ggaggctgag gcaggagaat tgcttgaacc cgggagatgg aggctgcagt
 gagctgaggt caggccactg cgctccaacc tgggcaacag agtaagactc catctcaaaa
 aaaaaaagc tggatttga tggaatatt aataacattc tccctctctc tccctttgcc
 25 tgtgtctcca tctctgtctt tttctgcatt tcttcatctc tgtactttcc atctctgtgt
 gtctgttccc atctgtctct ccatctatgg gcatctctgg gtctctcatg tctcctctg
 cccactttgc cacatctctg cctctctcat gcccccttt ctctcctgca ggggtgattct
 ggggggcctg tgggtctgcaa tggctccctg cagggactcg tgtcctgggg agattaccct
 30 tgtgcccggc ccaacagacc ggggtgtctac acgaacctct gcaagttcac caagtggatc
 caggaaaacca tccaggccaa ctccctgagtc atccaggac tcagcacacc ggcattccca
 cctgctgcag ggacagccct gacactcctt tcagaccctc attccttccc agagatgttg
 agaatgttca tctctccagc ccctgacccc atgtctcctg gactcagggt ctgcttcccc
 cacattgggc tgaccgtgtc tctctagtgt aaccctggga acaatttcca aaactgtcca
 35 gggcgggggt tgcgtctcaa tctccctggg gcactttcat cctcaagctc agggcccatc
 ccttctctgc agctctgacc caaatttagt ccagaaaata aactgagaag

SEQ ID NO. 6

hk6 amino acid

MKKLMVVLSLIAAWAEEQNKLHVHGGPCDKTSHPYQAALYTSGHLLCGGVLIHPLWVLTAAHCKKPNLQVFL
 GKHNLQRRESSQEQSSVVRVVIHPDYDAASHDQDIMLLRLARPAKLSELIQPLPLERDCSANTTSCHILGWG
 KTADGDFPDTIQCAYIHLVSREECEHAYPGQITQNMLCAGDEKYGKDSCQGDGGPLVCGDHLRGLVSWGNI
 PC GSKEKPGVYTNVCRYTNWIQKTIQAK

SEQ ID NO. 7

KLK6 nucleic acid

CDS 147.. 881

gtgcacccac gcgtccggct ggctggctcg ctctctcctg gggacacaga ggtcggcagg
 50 cagcacacag agggacctac gggcagctgt tccctccccc gactcaagaa tccccggagg
 cccggaggcc tgcagcagga gcggccatga agaagctgat ggtggtgctg agtctgattg
 ctgcagcctg gccagaggag cagaataagt tgggtgatgg cggaccctgc gacaagacat
 ctcaccccta ccaagctgcc ctctacacct cggggccactt gctctgtggt ggggtcctta
 55 tocatccact gtgggtcctc acagctgccc actgcaaaaa accgaatctt cagggtcttcc
 tggggaagca taaccttcgg caaagggaga gttccagga gcagagttct gttgtccggg
 ctgtgatcca ccctgactat gatgccgcca gccatgacca ggacatcatg ctgttgccgc
 tggcacgccc agccaaactc tctgaactca tccagccctt tcccctggag agggactgct

```

cagccaacac caccagctgc cacatcctgg gctggggcaa gacagcagat ggtgatttcc
ctgacacccat ccagtgtgca tacatccacc tgggtgtccc tgaggagtgt gagcatgcct
accctggcca gatcacccag aacatgttgt gtgctgggga tgagaagtac ggggaaggatt
cctggccaggg tgattctggg ggtccgctgg tatgtggaga ccacctccga ggccttgtgt
5 catgggggtaa catcccctgt ggatcaaagg agaagccagg agtctacacc aacgtctgca
gatacacgaa ctggatccaa aaaaccattc aggccaaagt accctgacat gtgacatcta
cctcccgacc taccacccca ctggctgggt ccagaacgtc tctcacctag accttgccctc
ccctcctctc ctgcccagct ctgaccctga tgcttaataa acgcagcgac gtgaggggtcc
tgattctccc tggttttacc ccagctccat ccttgcata ctggggagga cgtgatgagt
10 gaggacttgg gtccctcggtc ttaccccccac cactaagaga atacaggaaa atcccttcta
ggcatctcct ctccccaacc cttccacacg tttgatttct tccctgcagag gccagccac
gtgtctggaa tcccagctcc gctgcttact gtcggtgtcc ccttgggatg tacctttctt
cactgcagat ttctcacctg taagatgaag ataaggatga tacagtctcc ataaggcagt
ggctgttggg aagattttaag gtttcacacc tatgacatac atggaatagc acctgggcca
15 ccatgcactc aataaagaat gaattttatt atgaaaaaaa aaaaaaaaaa aaaaaaaaaa
agggcgccgc c

```

SEQ ID NO. 8

KLK6 nucleic acid

20 mRNA join(2001..2185,3084..3135,3559..3606,4346..4502,
8122..8369,9791..9927,11805..12483)

CDS join (3567..3606,4346..4502,8122..8369,9791..9927, 11805..11957)

```

25 acacttaaaa aatcttctga cttaaaaaaa aaagtatggg gatttgaaaa tgtaaagtgt
catgcgtgct tggcatcaca ttctattggc caggacttcc ctggatgcta aaggctcctca
aatgccaggc tggggggctg ggacttggtc ccaagggaga tggggaccca gggcacgtct
gtgagaggag gggcaaggct agcacaaggc acaggaaggc ctctctgggg caagggatac
agagaacaga gggatcctgg tccaggtggg agaggtgcag ctctgagttg ggggtgaggg
30 tgtgggtaca gagaggaagg gacccccag agagaggagg cagagggata gggcctgggtc
actgggttgt gcaacatcag acttgctgtc tgtgaagata gcacgtcctg agaagaaggc
gctgagggtca gtggggacca aatgtgagag ggagcaccgc gagagtatac tgaataccga
agtagtcttc atccctggag tgatgggggg tgcacaatgc aagatgacaa ttagattcaa
tgcaagacaa agaaaagggt tggctgggaa cagtggctca tgcctatggt cccagctcct
35 ggggaagactg aggcgggagg gtcgcttgag ccaggaggg ttgaggctgc cagagcaag
gatcgtgcca ctgactcca gcctaggcga cagaacaaga ccttgtctca aaagaaaaaa
gaactttttt ttttaagtta cctgtagtgc ccagcccaag caggtgctga gccagacttc
attcctatca ttgtccttat tacgcagtga cttccccctc ctcatctctc tccactctgc
cagcacaca ccctcacctc ccagcccata ccaaccaccc caaccactgc ctgtggtttc
40 ccatgtgcac ccaggccagg cattttcact gcttttctc ctgacctacg cctggctcag
ctttctaggc ccaagttcaa agacacctcc ctaaactctc ccagatccct ctgctactgc
ccagcaccac catcttatca cagccccacg tcgttcccaa gtgctctccg atttctgctt
aactccatgc ctctcgctgt gtgtccgcac ctcatcaata agtctcaag tctcttcca
tctgtctagc ttctcatcgc ctcggaatc atccccgcta cttcctgggg aaactgactc
45 ccttctgggc acacacagtg ctacccccgg ggaatctaa gaagagaccc aggagaagat
aagcacggag agtcagagaa tcaaggggaa agaaaggag agaggccggg cacagtggct
cacacctgta atccagcact ttggggaggc aaggtgggtg gatcacctga ggtcaggagt
ttgagaccag cctggccaac atgggtgaaac ctcttccta ctaaaaatac aaaaaacatt
tagccgggcg tgggtgtggg tgctgtaat ccagctact tgggaagctg aggcaggaga
50 actgcttgaa ctcaggaggc ggaggttgca gtgaactgag atcacaccac tgcactccag
cctgagtgc agagcaagac tccgtcaaaa aaaaagaaaag aaagaaagaa aagaaggaag
gaaagaaaga aggaaggaaag gaagggagga agggagagag gaagggagag aggaaggagag
agagagaaaa aaagagggag agagacacaa atacagagac tgagatggga gagagagaga
gatggaagct ccctcccctc catggccagg gagacagatg gagcaagaga cctcaggggt
55 gggcagactt ggaggagaag gaccaggagg atgtggagtg ccgaaatctc cagtcagggc
caggtgggca gtcagagact gcaaaggagg actgtcagac agggacaaaa ggaagccatt
gatgtaaccg ccctcccggc tgcccggcgg aagagagggt gagggccggg ctgctgggag
catggcactg ggggtgctggg aggcggacaa agcccagattg ttctggggc ctttcccat

```

	cgcgcctggg	cctgctcccc	agccccgggg	aggggcgggg	gccagtgtgg	tgacacacgc
	tgtagctgtc	cccccggtctg	gctggctcgc	tctctcctgg	ggacacagag	gtcggcaggc
	agcacacaga	gggacctacg	ggcaggtgtg	tgagtacccc	caaccgcact	gaacctgggc
	aggctgcttc	ccagtgcggg	agggctctag	agccccggag	gagggcctgc	aggtccctgg
5	gtggcacaga	gagtgcctggg	ggtgcaggga	ggcctggggc	accatctgct	tgccccagag
	gccggaatth	gtcttcagac	actttctttc	tccaaaaccc	ggaggtctaa	ggactgagcc
	gactagaact	tcctctgcct	cagattcagg	ccccagcccc	tcctccctca	gacccaggag
	tttaggtcct	agccccctcct	ccctcagacc	caggagtcca	agttcccacc	tcctccctca
	gactcaggag	tccaggcccc	cagccccctcc	tcctcagac	ccaggagtcc	aagttctcac
10	ctcctccctc	agaccagga	gtccaggccc	caagccccctc	ctccctcaga	cgcaagggtc
	caggccccca	gccccctctc	cctcagactc	aggagtccag	gcccccaagc	ccctccctcc
	tcagaccag	gagtccaggc	cctcactgca	ctcagggacc	agtgtccctc	ccctggagg
	cctggctcagg	ggtcaccaag	agcagagcgt	gggggaggga	ggaatgtgtg	tgaggaggcct
	gggtaaggag	gaaaagggtg	tagccagtct	cctggctcag	ggacctgaga	gacaggggtt
15	aaaaggacgt	tccagaagca	tctggggaca	gaaccagcct	cttccaggga	ggcctgggag
	ctgggggtgt	gtgtctggca	gtccctgcag	ccctgggctc	tgccggccct	gcgtccctcc
	cttggtctctg	ccactgcctc	tgagtgtctt	ctctctctac	ggctccccgc	atttctaaat
	ctttctgcct	cctcgtctca	aagctgttcc	ttcccccgac	tcaagaatcc	ccggaggccc
	ggaggcctgc	agcaggtgag	atcacagaca	tcacagaacc	tgccgggtgg	gcgggggtgg
20	tggtccattgc	gcacagagcc	aggctccgag	gaaaactccc	atacagagga	agaacgctag
	ggccccctat	ggtaaccctc	tcctgtcgac	aggaaggcaa	atcagtgcct	aagaaagtag
	aaagatctaa	tcagaatctc	accatgggtt	actggaccag	tgagcgtagt	tgaattctct
	ttggcactgt	tttcgtggat	cctcttgga	gatgtgggct	gaggaagaat	aaatcaggag
	gctagatggg	aaggacagag	gtcaaggcag	gagaccatag	caggccagga	aggaaggaga
25	ggatgcagag	ggagcagaca	gagggatggg	gggagggtcg	aggcagtgac	taatggacca
	tgtggcttcc	cctctcagga	gcggccatga	agaagctgat	ggtgggtgctg	agtctgattg
	ctgcagggtg	ggaaagggca	tttggtatgg	ggaggcttgc	agacagggtt	gggcttgttg
	atggagaaga	ggctggtatt	ggggatgggg	atatgcacag	ggttgggggtg	ggggagcttt
	gaaatgagga	agacgttggg	gattaggcta	aggggtggga	atacagatag	ggagggtggt
30	gggagggtgg	tttgaagata	tgagggtttg	gggtgggggt	ggctttaggg	atggggatct
	aaacatagaa	gaggtaggag	gtaggttgga	aagttggaga	gagcccgga	ataggggata
	cagttgggtt	tgtaatggga	atgggtaag	tttgggagtg	gaaatacaga	gaagcttttt
	ttttttgaga	cagggtctca	ctctgtcacc	caggctggag	tgtagtggca	tgatccatag
	ttcactgcag	acttgaactc	ttgggtctca	agtgaccctc	ccacctcagc	ctcccaagta
35	gctgggacta	caggcgtatg	ccaccatacc	ctgctaattt	gtgtgtgtgt	gtgtgtgtgt
	gtgtgtgtgt	gtgtgtgtgg	agatgaggtc	tcactgtgtt	accgaggctg	gtctcaaat
	cctgggctca	agcgatcctc	ctgcctcagc	tggtgattaca	ggcataagcc	actgcacctg
	accaatcttg	actggagttc	atgttgaggg	ggatgcgctt	ggtttctcca	gaactcctct
	ctgactcaga	tcctctctcc	ctcagcctgg	gcagaggagc	agaataagtt	ggtgcattgg
40	ggaccctgcg	acaagacatc	tcacccctac	caagctgccc	tctacacctc	ggggcaactt
	ctctgtgggt	gggtccttat	ccatccactg	ttgggtcctca	cagctgcccc	ctgcaaaaaa
	ccgtgagtct	acactgtaaa	tgaacagcag	atgcgactga	accctgaggg	tgtcttatag
	atgtcaggca	ggaggtgaca	taggcattcc	ccccatccca	gcacgaggcc	atctgatagc
	caggtgcatt	cggctgttgc	ttaattgagt	acttaattgt	tgccaggccc	tgccggcata
45	gcagtggaaa	agaaaataaa	aaaaagaaaa	caaaaaaaaa	caagcaaaat	tgctgttttc
	ctgaacttac	tttctaattg	gggaattgga	tcattttggg	acctgcaggg	cgtgatgggc
	atgttgattt	aattctgagc	acagtaggaa	ggcactgggc	agttttgttt	ttgttgtttg
	tttgtttttt	gagacacagt	ctcgctctgt	cacccaggct	ggagtgtagt	ggcatgatct
	cagctcactg	caacctctgc	ctcccagggt	ccagcgattc	tcctgcctca	gcaccccaag
50	tagctgagat	tacaggtgtg	caccaccttg	cctggctaata	ttttgtatgt	ttggtagaga
	cggggtttca	ccatgttggc	caggctgggtc	tcgaactcct	gacctcaggt	gatccgcccg
	cctcgccctc	ccaaagagct	gggattacag	gcatgagcca	ccaccacacc	cagcctgatt
	tacattttta	caagcaccct	ggctaccacg	tggaacgtgg	tctgggcaag	agagagggag
	ggaggcccac	gtgggggtg	ttgctttctc	ccggcgacat	aggaggggtg	cttgaaacca
55	ggcggctcgca	gtggggatgg	agggatgttg	aatatctttg	gatgtggaat	tctgagactg
	agccagcaga	atctggcaac	gaggaacagg	agggagagga	agaagcacgg	ctggcttccg
	tgtatttgtc	ctgaacaact	gggtgttttg	ccacgtcttt	ctctgagttg	tgggagaggg
	aaagagaaac	aggccgggtg	taggcagggg	agcatctgac	atthtgcctt	agccacgatg
	agttggagat	gccggggaga	tgtcccagca	gggaggccag	ggaggactct	ggagctcaga

	ggagagggtca	gggctggagg	ttaaaatgaa	ggcatcgtca	gcaaacaggt	gtattttaaag
	ccatgggact	agatgagatc	atccaaaaag	ctggcatagt	tggaggagct	ggagggccca
	ggacaaaaac	cctgggcgct	gatcctcact	agtcagattc	acgacagctg	ccacttgttt
	gatgctaact	accaatcagg	tgctgagtga	aaccatgtac	acacctttcc	tggaaatgcc
5	accacaaggg	actcctggca	ccattttgca	aatgaggaaa	ctgagggtga	gggaaatagc
	aagtgcacaat	ccctgggggtg	gttccctga	ccccaaaggag	accttggatg	actctacca
	ccatcattca	ttcctttgat	gtacattgac	taagagcacc	tgctaagtgc	cacattcgag
	ttgggcagtg	gagattcagc	aatggatggg	acacacacgt	catccctgcc	ctcgggagca
	caaggacaga	aaggtgcaga	caagcaaagt	gagggctggg	catggtggct	cacgcctgta
10	atcccagcac	tttgggaggg	cgaggtgggt	ggattacctg	agttcgagac	cagcttggcc
	aacatggctc	aacctgtct	ctactgaaa	tacaaaaaat	tagccaggcg	tgggtggggg
	cttctgtaat	tccagcaact	tgggaggcta	aggcaggaga	attgcttgaa	cgtgggaggg
	ggaggttgca	gtgagccgag	atcgcgccac	tgactccag	cctgaaccac	agagcgagac
	tctgtctaaa	aaaaaaaaaa	ggaaagaaag	aagcagcaaa	ttgggctggc	cgtgggtggc
15	catgcctgta	atcccagcac	tttgggaggg	cgaggcgggt	ggatcactcg	agcccaggag
	tacaaagctg	cagtgcagctg	tgatctacag	aacaccactg	cagatccagc	ctgggtgaca
	gagcgagacc	ctgtctcaaa	aaaacaaaca	aacaaaagaa	gcaaaccctt	caaaacccca
	tataattaca	aattatgaag	gaaaagaata	cgggtacctg	ctttagatgg	aggagggtca
	ggaaggactt	tctaatagaga	taaaatccaa	gcggaggcat	gaagatggga	aaaggaatgt
20	tcagggcaga	ggaaaggctg	tgataacacc	cctgaggtga	gaaccgtctt	gagtattctc
	agaaaaataaa	atttcccggt	cactgggggg	cagaaggtgc	tgggagataa	ggttggaaag
	tgactacagc	cagatcacac	aggggctcca	gtgccaagtg	gaggagccca	ggctttattc
	ttaggacaat	ggggagccat	gggtgatgtc	tgagcaaggg	agtgactctc	tgtttcagga
	atatgtatca	aacacctatc	ctgtgccagg	tgctgatcaa	cgcactggag	atactatata
25	tgaatagaac	aaaaatcccc	atcttgacat	cctagagctg	cactgtctaa	tatggtagcc
	atcagccaca	tatagcaaat	tacattgaaa	ttaatgaaat	ggaaaatcca	caagccacat
	ttcaagtact	cagcagccac	ctgtagcttg	tggttcccc	agccacctct	ggacagtgca
	gatcgagatc	atggcatcgt	agcatttagt	ggacagcatt	gctctgcaag	gaggagaaat
	aacacaattga	gtaaatat	aacaataaat	atatagcagg	tcggatgatt	gtgataggtt
30	ctctggtgga	acagaaagca	ggggaggagg	ataggaaattg	cctactaaca	ggtatttgta
	ttttaattgg	gcaactaagg	aaggcttccc	tgagaggcga	catttaaagg	aagtggaggga
	gtgagctatg	cagatacttg	gaggacagac	ttgctggcag	agggaaacagc	agtgcaaagg
	ccctgaggtg	ggaagatcac	tattgtgttc	aaggcaagac	aggggaagcca	gogtttggct
	ggagcagagg	gagagaaggg	gagagtggga	ggagaagatg	tctgtgagat	gatggggcag
35	tgcttgcaag	gcctggtgtg	ccacgttgag	aactttggct	ttgattctga	gtgagatggg
	agtcatagga	ggggctgagc	agaggaggca	caggaccaac	ttacattggt	aaaatatctc
	tggttgcttt	gtggaggatg	gactgtgggg	gaccagagac	agagcaggga	gcccagtgag
	gaggctactg	ctctagtcca	ggtaggaagt	gaaaaggcag	ctcaaaccac	gatggtagcc
	gtgggaaagg	tgagatgtgg	ccagattctg	gatatgcttc	agagaggcaa	aaggaattct
40	ggacagcttg	gatgtagggc	atgaaataaa	gagagtgaag	aatagccccc	aagattattc
	tgaaaggatg	gaattgccat	ttaccagct	ggggaagact	gtgggaggag	caggccagcg
	attcatgact	tcccagccct	ctctgaagcc	tcaactgcag	cccaagggtc	ccaggtgaga
	cccagccctc	ttccttccca	ggaatcttca	ggtcttctcg	gggaagcata	accttcggca
	aagggagagt	tcccaggagc	agagttctgt	tgtccgggct	gtgatccacc	ctgactatga
45	tgccgccagc	catgaccagg	acatcatgct	gttgcgccctg	gcacgcccag	ccaaactctc
	tgaactcatc	cagcccttcc	ccctggagag	ggactgctca	gccaacacca	ccagctgcca
	catcctgggc	tggggcaaga	cagcagatgg	tcagtagtgg	gaggctgggtg	gggagcaggc
	tactggctac	ttggggaagt	gtgccaaagg	atggggagtg	ggaaaattgg	tgaggggcca
	tgggaagatg	ggctaagtgt	gaggaccaat	gggacagggg	ttcaatggga	gaaaggtcaa
50	gggggaggga	gagtgaattt	gggagctggg	ccagtgcagt	aacagccaat	ggaaaatgta
	gaccaatggg	tgaatagcat	gggagagatg	gaacataaga	tgaaggttca	ataaagaggg
	aagggtcagt	gggagatgct	aatcaggaag	gatgtcaaa	gtcaaagggg	actgatcagg
	attcatgaa	cagcaggaag	gaataatgga	gaaggaaact	atggaagaag	agaaaccaat
	aaagcacaaa	agccaactga	aggatgtgaa	ttgagacagt	gaatgggggt	atagtgtatg
55	gaagagggac	taaggggaaa	ggatcaatgg	tcagagaggag	tcactagagg	aaaaaacagg
	tccaatagat	cagcaggatc	catgaagggtg	ggcctgtgtg	tgaagggcca	ataagaaagg
	tgaaccattg	gatgaagggc	cagtgggaag	gcagagacaa	tgggggaggga	tgccgcaagt
	tagaaaagga	ccaatgaggg	aggtggacca	ttggatgaag	ggctaataag	aagggagagc
	cagtggggga	tggtagggcc	agttagaaaa	ggaccaagga	gggaagcaga	ccaataggaa

	gagagagcca	atgagggagg	gcagggccag	ttaggaaagg	accaatgagg	aaggtagacc
	attggaggaa	gggccaatag	aaagggagga	tccatgaggg	aggggtggga	cagttagaaa
	aggaccaatg	atggaggtgg	accattggat	gaagaacca	tagaaaggaa	gaaccaatgg
	gagagggcat	ggccagttag	gaaaagacca	atggtcacag	agtgaccaat	caagatgaat
5	caatgggcag	gaagtgtoca	atgaagaatg	gactactgat	caggaggggt	acagtagagg
	agggcgtaac	agaggaagag	tcctccaggt	caactgaaac	tactgaagaa	gggtgggacca
	gtggaagaga	gaaaagtgga	ggagggacct	aagagaaaag	gaaaaccaat	aggaaatgag
	gactcctgga	gaagagacta	ttaatgagga	agacagccaa	tgggggggaa	gaatgataga
	aagagggacc	aattaggagg	cagggacgat	ggtaatgaga	tgtaaagaatg	agagacaaac
10	aggaagaggg	gtgccaatag	aaaagaggga	ccaatagagg	atggaggact	tataggggtt
	gggggggtgac	tggggaggat	gaggggagtg	caaggcctgg	gctgagtctg	gcccactctt
	cccctaacag	gtgatttccc	tgacaccatc	cagtgtgcat	acatccacct	gggtgtcccgt
	gaggagtgtg	agcatgccta	ccctggccag	atcacccaga	acatgttgtg	tgctggggat
	gagaagtacg	ggaaggattc	ctgccaggtg	aggtgaccgg	gatctgccac	ttacacagcc
15	agggacagga	cgaagtcaca	aaaacatggc	cagacacagg	aagagagaga	cacaggccaa
	aagagagctt	tacagagaca	gatagagaca	ggctgaggga	gaaccaagc	cttgaagaag
	agagacttag	ttcaacacac	agagacacag	tcagggatat	gcagagatat	aagacacag
	ccagcagaga	caggaagtgc	agagacaagg	atggaggccg	cgggatcaag	aaccagagag
	gccaggagca	gcggctcatg	cctgtaatcc	cggcactttg	ggaggccgaa	gcaggaggat
20	cacctagggg	caggagttcg	agaccagcct	gatcaacatg	gtgaaaccct	atctctacta
	aaaatacaaa	aattaggatg	ggcacagtgg	ctcatgcctg	taatcccagc	accttgggag
	gccgaagcag	gaggatcacc	tggggtcagg	agttcgagac	cagcctgatc	aacatgggtg
	aaccctatct	ctactaaaaa	tacaaaaatt	aggatgggca	cagtggctoa	tgctgtaat
	cccagcacct	tgggaggccg	aagcaggagg	atcacctggg	gtcaagagat	tgagaccagc
25	ctggccgata	tgggtgaaacc	ctatctctac	taaaaaataca	aaaattagct	gggcctgggtg
	caggcgccctg	tagtcccagc	tactcaggag	gctgtggcag	gagaatcact	tgaacctgga
	ggcggagggtt	gttgacagtga	gtcgagatca	tgctactgca	ctccagcctg	gcaacagagc
	aagattccgt	ctcaaaaaaa	aaccaaaaaa	caaaaaattac	gcaagcatgg	tgggacacac
	ctgtagtccc	agctactcgg	gaggctgagg	ctggagaatt	gcttaaacc	aggaggcaga
30	ggctgcagtg	agctgagatc	acgccactgc	actccagcct	ggggacagag	ccagactctg
	tctaaaaaca	aaaagaacca	aagagaagta	gtaagggaagc	agatgggtgtg	aggggactgt
	ccttctctcaa	acagagcccc	cacgagtcct	gctcagaaac	gaccaggctc	tggaggaggg
	agacactagc	tggggaaaagg	ggactccctc	ccgaatactt	taacttgggt	ttcctccatt
	gtcatccatc	caggctctcc	tctttatgcc	agaatgacta	atgcaactgag	ggatgtgcag
35	agaccaacca	agggggagac	acaggcagaa	acggagacac	aggcagaaac	agggacagag
	acagggaaag	cgatacatag	caagtgggac	gcaaagaaaag	ggcagggtggg	cgagactgtc
	ctcaagacac	gaggtggaga	ggtgtccctg	gacagaatag	tgccaggcat	atctctccct
	gggcccctccc	tacctctccc	acctgggtct	tatcgtctcc	tctcccccct	cctccctctc
	ctcctcttcc	tctctctcct	cctccctcat	catcttcttc	ttttctctct	ctctccatcg
40	gtctctacac	ctctgcctct	ctccacacct	ctcagttctc	attctttaa	tgtttctctt
	tcttgctctc	tatgttctct	tgcatcttgg	cattcctatc	tctgtgtctt	tgagtctcct
	ttattctctc	tctaccatct	tctctctgtg	cctttgtgtg	tcttactgtc	tctctctctg
	tctctctgtc	cctgagtctt	tctctccatc	tttcagtaag	tacctctgtc	cctttctacc
	tctctctctg	tcacacacac	acacacacac	acacacacac	acacacacac	acacacagtc
45	tctgggtttc	tatctgtatc	tgactttctc	cctctttcct	gcagggtgat	tctgggggtc
	cgctggtatg	tggagaccac	ctccgaggcc	ttgtgtcatg	gggtaacatc	ccctgtggat
	caaaggagaa	gccaggagtc	tacaccaacg	tctgcagata	cacgaactgg	atccaaaaaa
	ccatttaggc	caagtgacct	tgacatgtga	catctacctc	cogaactacc	accccactgg
	ctgggtccag	aacgtctctc	acctagacct	tgccctccct	cctctctctg	ccagctctga
50	cctgatgct	taataaacgc	agcgacgtga	gggtcctgat	tctccctggg	tttaccctag
	ctccatcctt	gcatacactg	ggaggacgtg	atgagtgagg	acttgggtcc	tgggtcttac
	ccccaccact	aagagaatac	aggaaaatcc	cttctaggca	tctcctctcc	ccaacccttc
	cacagtttg	atttcttctt	gcagaggccc	agccacgtgt	ctggaatccc	agctccgctg
	cttactgtcg	gtgtccctct	gggatgtacc	tttcttcaact	gcagatttct	ccactgttaag
55	atgaagataa	ggatgatata	gtctccataa	ggcagtggtc	gttggaagaa	tttaagggtt
	cacacctatg	acatacatgg	aatagcacct	gggccaccat	gcactcaata	aagaatgaat
	ttt					

SEQ ID NO. 9

KLK6 nucleic acid

5 CDS 246..980

```

aggcggacaa agcccgattg ttcttggggc ctttcccat cgcgcctggg cctgctcccc
agccccgggc agggggcgggg gccagtgtgg tgacacacgc tgtagctgtc tccccggctg
gctggctcgc tctctcctgg ggacacagag gtccggcaggc agcacacaga gggacctacg
10 ggcagctgtt ccttcccccg actcaagaat ccccgaggc cccgaggcct gcagcaggag
cggccatgaa gaagctgatg gtggtgctga gtctgattgc tgcagcctgg gcagaggagc
agaataagtt ggtgcatggc ggaccctgcg acaagacatc tcaccctac caagctgccc
tctacacctc gggccacttg ctctgtggtg gggtccttat ccatccactg tgggtcctca
cagctgcccc ctgcaaaaaa ccgaatcttc aggtcttcct ggggaagcat aaccttcggc
15 aaagggagag ttcccaggag cagagtcttg ttgtccgggc tgtgatccac cctgactatg
atgccgccag ccattgaccag gacatcatgc tgttgccctt ggcacgccc gccaaactct
ctgaactcat ccagcccctt cccctggaga gggactgctc agccaacacc accagctgcc
acatcctggg ctggggcaag acagcagatg gtgatttccc tgacaccatc cagtgtgcat
acatccacct ggtgtcccggt gaggagtgtg agcatgccta ccctggccag atcaccacga
20 acatgttgtg tgctggggat gagaagtacg ggaaggattc ctgccagggt gattctgggg
gtccgctggt atgtggagac cacctccgag gccttgtgtc atggggtaac atccccgtg
gatcaaagga gaagccagga gtctacacca acgtctgcag atacacgaac tggatccaaa
aaaccattca ggccaagtga ccctgacatg tgacatctac ctcccgacct accacccacc
tggttggttc cagaacgtct ctcacctaga ccttgccctc cctcctctcc tgcccagctc
25 tgaccctgat gcttaataaa cgcagcgacg tgagggtcct gattctccct ggttttacct
cagctccatc cttgcatcac tggggaggac gtgatgagtg aggacttggg tctcgggtct
tacccccacc actaagagaa tacaggaaaa tcccttctag gcatctcctc tccccaaacc
ttccacacgt ttgatttctt cctgcagagg cccagccacg tgtctggaat cccagctccg
ctgcttactg tcgggtgtccc cttgggatgt acctttcttc actgcagatt tctcacctgt
30 aagatgaaga taaggatgat acagtctcca tcaggcagtg gctgttgga agatttaaga
tttcacacct atgacataca tgggatagca cctgggccgc catgcactca ataaagaatg
tatttt

```

35

SEQ ID NO. 10

hk7 amino acid

```

MARSLLLPLQIILLLSIALETAGEEAAQGDKIIDGAPCARGSHPWQ
40 VALLSGNQLHCGGVLVNERWVLTAAHCKMNEYTVHLGSDTLGDRRAQRIKASKSFRHP
GYSTQTHVNDLMLVKLNSQARLSSMVKKVRLPSRCEPPGTTCTVSGWGTTTSPDVTFP
SDLMCVDVKLISPQDCTKVYKDLENSMLCAGIPDSKKNACNGDSGGPLVCRGTLQGL
VSWGTFPCGQPNDFGVYTQVCKFTKWINDTMKKHR

```

45

SEQ ID NO. 11

KLK7 nucleic acid

CDS 16..777

```

ggattttccgg gctccatggc aagatccctt ctcttgcccc tgcagatcct actgctatcc
50 ttagccttgg aaactgcagg agaagaagcc cagggtgaca agattattga tggcgcccca
tgtgcaagag gctcccaccc atggcagggt gccctgctca gtggcaatca gctccactgc
ggaggcgctc tggtaaatga gcgctgggtg ctcaactgcc cccactgcaa gatgaatgag
tacaccgtgc acctgggcag tgatacgctg ggcgacagga gagctcagag gatcaaggcc
55 togaagtcac tccgccaccc cggctactcc acacagaccc atgttaatga cctcatgctc

```

```

gtgaagctca atagccaggc caggctgtca tccatggtga agaaagtcag gctgccctcc
cgctgcgaac cccctggaac cacctgtact gtctccggct ggggcactac cactagccca
gatgtgacct ttccctctga cctcatgtgc gtggatgtca agctcatctc ccccaggac
tgcacgaagg tttacaagga cttactggaa aattccatgc tgtgcgctgg catccccgac
5 tccaagaaaa acgcctgcaa tggtgactca gggggaccgt tgggtgtcag aggtaccctg
caaggtctgg tgtcctgggg aactttccct tgcggccaac ccaatgacct aggagtctac
actcaagtgt gcaagttcac caagtggata aatgacacca tgaaaaagca tcgctaacgc
cacactgagt taattaactg tgtgcttcca acagaaaatg cacaggagtg aggacgccga
tgacctatga agtcaaattt gactttacct ttccctcaaag atatatttaa acctcatgcc
10 ctgttgataa accaatcaaa ttggtaaaga cctaaaacca aaacaaataa agaaacacaa
aaccctcaa

```

SEQ ID NO. 12

KLK7 nucleic acid

15

mRNA

```

join(1756..1785,3179..3309,3722..3869,4566..4813,5129..5265,7362..8265)
/product="stratum corneum chymotryptic enzyme" /note="alternatively
spliced"

```

20

```

mRNA join(1756..1785,3179..3309,3722..3869,4566..4813,
5129..5265,7362..7991) /note="alternatively spliced"

```

25

mRNA

```

join(1821..1864,3179..3309,3722..3869,4566..4813,5129..5265,7362..8265)
/product="stratum corneum chymotryptic enzyme" /note="alternatively
spliced"

```

30

mRNA

```

join(1821..1864,3179..3309,3722..3869,4566..4813,5129..5265,7362..7991)
/note="alternatively spliced"

```

```

CDS join(3237..3309,3722..3869,4566..4813,5129..5265, 7362..7517)

```

35

```

ggcatggtgg tgcacgcctg taatccagct actcaggact ctgaggcagg agaatacactt
gaacacgggg gagtggagggt tgcagtgagc cgagatcgtg ccattgcact ccagcctggg
tgacagagcc agagtccatc aaaaaaaaaa aaaaataaga aagattcttc tctcctctat
gtgtccatgc agtctcatca ttttagctacc acttgtaagt aggaacatgc catatctggt
40 tttctgttcc tgccttagtt tgtaagggtg atggcctcca gctccattca cgtccctaca
aaggacatga tcgtgttctt ttttatggct acgtagtatt caattgtgta taagtaccac
atcttcttaa tccagtctat cactgatgga catttagggt gattccctgt gtttgcgtgt
gtcaatagtt ctacaatgaa cgtacgtgtc catgtgtctt taaacagaat gatttatatt
cctttgggta cacacactgg ggcttatgag aggggtggaga gtgggaggaa ggagaggatc
45 agaaaaaaaa aactaatggg tactaggctt aatacctggg tgattaaata atctgtataa
caaaccccca tggcgcaagt tcacctacgc aacaaacctg cacatcctgc acatgtacct
ccgaactgaa aagttaaaaa agaaaaaata aatatttgct tataaattaa taaatgaagc
cctcaaaaaa gttctattag ataatgttaa gtacagacat ttttgttata aatacataat
atacaaaagaa atctatgtat aacatgatta aaatgaccat agaacatag atcctaaaca
50 tggcaaatat tagtgggggtg ggggttaggga aagcgttggt tttaacttac acctctctgt
tagagttggg aatgggttca ggcgtaatta caggcacgac tgggatcagc ttggacaagt
tcccccaggc gggccagaat taggatgtag ggtctaggcc acccctgaga ggggggtgagg
gcaagaaaaa ggcgccagaa gccgggcgca gtggctcacg cctgtaatcc cagcactttg
cggggcccag gcgggcacat catgaggtca ggagatcgag accattcttg ccaacatagt
55 gaaacccggt ctctactaaa aatacaaaaa ttagctggga gtggtggtgc gtgcctgtaa
tcccagggtac tcgggagggt gaggcaggag aatcacttga acctgggagg cggagctggc
agttagccga gatcgcgcca ccgcactcca gcctggcgat agagagagac tocatccaaa
aaaaagaaag gaagggagggt agggaggagg gaagaaagaa agaaaaccgc ccagagagaag

```

	gacccgagcc	agagcctatt	ctctgagctc	agcgactgct	tgaatcccgc	tcctgcccct
	cagacccagc	gcacccgggtc	cctcccccca	gagcagccag	gagggactgt	gggaccagaa
	tgtgcggggg	cgcaggagct	gggcaccgcc	cgtccctcgg	agggagggtg	gagagagagt
	gcagtgggtg	caattgctct	cgctgcgtca	gggttccaga	taaccagaa	cgcaaatgca
5	ggcgggggtg	tcccagagtc	ggctccgcct	gcaccccagg	gcgctggggc	cgggcatggg
	gcgggggggtg	atataagagg	acggcccagc	agagggatga	agatttttga	gccagctgt
	gtgccagccc	aagtccgaac	ttggatcaca	tcagatcctc	tcgaggtgag	aagaggcttc
	atcaaggggtg	cacctgtagg	ggagaggggtg	atgctggctc	caagcctgac	tctgctctcg
	agaggtaggg	gctgcagcct	agactcccgg	tcctgagcag	tgagggcctg	gaagtctgca
10	atlttggggcc	tttttagggaa	aaacgaacta	cagagtcaga	agtttgggtt	ccacagggaa
	gggcaagatc	ggagcctaga	ttcctgggtc	tcagggatc	tgaagaacag	gaatttttggg
	tctgagggag	gaggggctgg	ggttctggac	tcctgggtct	gagggaggag	ggcctggggg
	cctggactcc	tgggtctgag	ggaggagggg	ctgggggtct	cgactcctgg	gtctgagggg
	ggaggggctg	ggggcctgga	ctcctgggtc	tgagggagga	ggggctggga	cctggactcc
15	taggtctgag	ggaggaggag	ctggggcctg	gactcctggg	tctgagggag	gaggggctgg
	ggcctggact	cctgggtctg	agggaggagg	ggctggggcc	tggactcctg	ggtctgaggg
	aggaggggct	ggggcctgga	ctcctgggtc	tgagggagga	ggggctgggg	cctggactcc
	tgggtctgag	ggaggaggag	ctggggcctg	gactcctggg	cctgagggag	gagggactga
	gacctggact	cctaggtctg	agggaggagg	gactgggacc	tggactcctg	ggtctgaggg
20	aggaggagct	ggggcctggg	actcctgggt	ctgagggagg	aggggctggg	gcctggactc
	ctgggtctga	gggaggaggg	gttggggcct	ggactcctga	gcctgagggg	ggagggactt
	ggacctggac	tcctaggtct	gagggaggag	gagctggggg	cctggactcc	taggtctgag
	ggaggcgggg	ctgggggcct	ggactcctgg	gtctgagggg	ggaggggttg	gggcctggac
	tcttgagcct	gagggaggag	ggacttggac	ctggactcct	aggtctgagg	gagggaggag
25	tgggggcctg	gactcctagg	tctgagggag	gaggggctgg	gggcctggac	tcctgggtct
	gagggaggaa	ggtgctaggg	tctggactct	tgggtatgag	ggaggaggag	gttaggggtc
	tggacttctg	agtgttaagg	aggagaggcc	agagaaagga	atlttctgggt	ctgagggagg
	aggggctggg	gttctggacc	cctaggtctg	agggaggagg	ggctggggcc	tggaccctctg
	ggtctgaggg	aggaggggct	ggggccggta	ctcctgggtc	tgtgggggga	ggggctgggg
30	cctggacccc	tgggtctgag	tggggagggg	ctgggcctga	atgctttctc	cttctcagct
	ccagcaggag	aggcccttcc	tcgcctggca	gcccctgagc	ggctcagcag	ggcaccatgg
	caagatccct	tctcctgccc	ctgcagatct	tactgctatc	cttagccttg	gaaactgcag
	gagaagaagg	tgaaagctgg	actgggaagt	ctgacctcac	ctcagggccc	ccactgaccc
	tctccaagga	gtccctgagt	cagaaccctt	ccctcctcaa	acagcttcca	tcctgggagg
35	accagactgt	cggtgaagc	ccccgctctt	cctgcttctg	ctgactcagg	gggtctctgt
	ccctccagct	ccctgocctc	tgtgctcagg	gtctctctgt	ggttccccag	atgagatgag
	cctcctgggt	ttctgagtg	gtccttcttg	tctgtctcta	tccttatctc	ttgttttctc
	tgtattttctc	cacacatttt	catctgtctc	tgtccatctc	tgactctggg	aatccctgag
	gtgcagcctc	agccttcccc	taatgctagc	tacccacgtg	ctcctccatg	tctccatcca
40	gcccagggtg	acaagattat	tgatggcgcc	ccatgtgcaa	gaggtcccca	cccatggcag
	gtggccctgc	tcagtggcaa	tcagctccac	tgcgaggcgc	tcctgggtcaa	tgagcgctgg
	gtgctcactg	cgcgccactg	caagatgaag	taggtgccgc	ccaagtctct	cctggagggtg
	caccagcgtc	tcagctcgcc	tatgggggtg	gaagggcagt	ctttctgtgc	ctaoggtctc
	attctcctct	ctctgggtct	ctgtcctcct	ctctctgggc	ctctgtaccc	cctctccctg
45	gggctctgtc	ccctctctct	cctggctctc	tgtctccctc	tctctgggtc	tctgtccccc
	tctctctgga	tctctgttcc	cctctctctg	tgtctctgtt	ccccattctc	tctaggtctc
	tgttccccct	cctctctctc	tgggtctctg	tcctctctct	tctggatctc	tgtccccctc
	tctctctggg	tctctgttcc	cctctctctg	ggctctctgt	ccctctctct	tctctgtgtc
	tctctccccc	ctcctctctc	tgtgtctctg	tcctccctcc	tatctctgtg	tctctctccc
50	ccctcctctc	tctgggtctc	tgtccccccc	tctctgggtc	tctgtctccc	tctctctggg
	gctctgtccc	cctctctctc	tggatctctg	ttccctctct	tctgggtctc	tgtctccccc
	cctctctctg	tgtctctgtc	ccccctcctc	tctctgggtc	tctgtcccca	ccccgtcccc
	caggctcttg	caacccctct	ctgtcacagt	gtctcttctg	aatctgtgaa	tgtcacctct
	cgcagtgagt	acaccgtgca	cctgggcagt	gatacgctgg	gcgacaggag	gctccagagg
55	atcaaggcct	cgaagtcatt	ccgccacccc	ggctactcca	cacagaccca	tgttaatgac
	ctcatgctcg	tgaagctcaa	tagccaggcc	aggctgtcat	ccatgggtgaa	gaaagtcagg
	ctgccctccc	gctgcgaacc	ccctggaacc	acctgtactg	tctccggctg	gggcactacc
	acgagcccag	atggtaggtg	gcctcagtga	cccaggagtg	caggccccag	ccctcctccc
	tcagaccccag	gagtccaggc	ccccagcccc	tcctccctca	gacccaggag	tcaggccctc

	agccccctcct	ccctcagacc	caggagtcca	ggccccccagc	ccctcctccc	tcagaccggg
	gagtccagac	cccagccct	cctccctcag	accagcagc	cctggggccc	agaccctcct
	ccctcggaac	caggagcctg	aacaacagcc	cttctgggtcc	tcgcccccat	cctctctgac
5	tgacagctct	ccctgctcct	ccctgcagtg	acctttccct	ctgacctcat	gtgogtggat
	gtcaagctca	tctccccca	ggactgcacg	aaggtttaca	aggacttact	ggaaaattcc
	atgctgtgcg	ctggcatccc	cgactccaag	aaaaacgcct	gcaatgtgag	accctcccc
	ccaattcctc	cccagtcctg	ggtaccctgt	ctgcatgccc	cagggacaga	gcttgaccgg
	agtgactggg	taccaagccc	ggccttgccc	tccccccagg	cctggcctcc	tcagcttttt
	ccacctcatt	ctctgcctag	gtcaggggtg	ggagtttact	taggggcca	tgtggccctg
10	gggatgagac	agagagttta	ataggggtga	gaaagtggg	gtgggaccag	ggaaggagac
	tgaggtgctg	gcctcagggc	caaaccttaa	gggggcacca	aaaacctcag	tgattgagat
	aaatcataat	gcaatattta	aaaataaaaa	taaaaactca	tgcagaagtc	catgatggac
	aaaatgtcac	attttaata	aagagcaggt	ggatcttact	gaattttccc	ttgcogtaag
	tactagcgtg	gctcagcaca	gcgctgtact	ggcactgtct	tcatttaaaa	tgtggatacc
15	atgcccata	tgcagtttta	tgtattacat	ttgatttcgt	taagtactgc	attgaagtat
	tgtgtattgc	agttactgag	attttgtgcc	tgaagctgat	gactcactca	cctgaccctg
	gcctgggtcc	cggggaaaac	actctttctc	tccacctcct	ctctgttccc	ctttcttgge
	cttttgcatt	ccctctgttt	tctgaacagt	cttcccat	ctctctttgt	gacataat
	catttcattc	ttttcctctt	tgttttttct	ctgtgttgag	ctagcttgct	ctccctccct
20	tgttctctct	ccatgccttc	ctctctgctc	tctgtcttct	ccctctttct	cttgccttct
	tctctctcct	ccctcccttc	tctcctctct	ctgccccct	gctctctctt	ttttcctctc
	tctctgtctc	ctctctggcc	ctctcctctt	tctctctctc	ccccacttct	ctgtctctct
	tcatctctct	ccctcatctc	tccctggccc	ctccttttta	ctgtctctct	ctttctcttt
	cttctatctc	tctcctctcc	ctgcctctcc	cccatctctg	tctttctttc	tctctcttta
25	ttctcctcct	ctcttccagt	ctctctctcc	tctccccacc	cccaccccat	ctctctcccc
	acaccttccc	cccttttctc	tttgtctctc	tcttctccct	ctttcttctc	cacccccatc
	tctctctctc	ttctcttccc	acacctctcc	catctccctc	atctctttgt	ctgtctctct
	tctccctcct	tcttttccac	ccccatctct	ctgtctctct	ctctccccat	accttttccc
	tcttctctat	ctctctttgt	ctctctctct	tttccctctt	tcttctccac	ctccaaactc
30	ctctgtctct	ccacacccat	cctccttgct	cacatctgca	ccttcagctg	tcaggggatg
	tgggatgggtg	agtgttaggg	atagaggaga	tgggagagag	atgactgtcc	tagagaatag
	ggtgttcccc	ttctccctctg	gtgaggggcca	gtttcatgaa	tgtgcaagct	ctgcaoggac
	acagagcccc	acactcagaa	gggtctcaaa	cttagtctaa	tgcattcctg	ctgttgtctt
	gaaattctca	ataatttttg	aacaaagggc	cctgcatttt	cgttttgcac	caagtctgt
35	aaattatgta	actggtcttc	accctggctc	ccgagaccat	cgtgtccccc	tttctctgce
	cacagggcac	gcattccacc	cttgagatg	atgttccctc	tccactagc	ttggagcagg
	gtccttaaca	ttggaaaata	aagagtgtct	tgatcctgga	agccccacc	cttctctgca
	attgggtctca	ttggccaagg	gtcaaaccag	tgtcttcaaa	ggacctagtg	tgtccctagc
	actagctctc	ccattagctc	ccagagacaa	tgagtctctt	ctcattgggt	atgggtggaag
40	tccataatct	gcaagacaaa	gaccgataac	tgaggaatgt	atgagaatga	gttgggcttt
	gatctgaagc	caaagttaat	ctccggctct	attccctcta	gggtgactca	gggggaccgt
	tgggtgtcag	aggtaacctg	caaggtctgg	tgtcctgggg	aaactttccc	tggggccaac
	ccaatgaccc	aggagtctac	actcaagtgt	gcaagttcac	caagtggata	aatgacacca
	tgaaaaagca	togctaacgc	cacactgagt	taattaactg	tgtgcttcca	acagaaaatg
45	cacaggagt	aggagcgcga	tgacctatga	agtcaaattt	gactttacct	ttcctcaaa
	atataatttaa	accaacctca	tgccctgctt	ataaaccaat	caaattggta	aagacctaaa
	acaaaaacaa	ataaagaaac	acaaaacctc	cagtgtctga	gaagagtcag	tgagaccagc
	actctcaaac	actggaactg	gacgttcgta	cagtctttac	ggaagacact	tggtoaacgt
	acaccgagac	cottatttcac	cacctttgac	ccagtaactc	taatcttagg	aagaacctac
50	tgaacaaaa	aaaatccaaa	atgtagaaca	agacttgaat	ttacctgat	attattttatc
	acagaaatga	agtgaaaoca	tcaaacatgt	tccaaaagta	ccagatgggt	taaataatag
	tctggcttg	cacaacgatg	ttttttttct	ttgagacaga	gtctctgttg	cttgggctgc
	aatgcagtga	tgcaatcttg	gtcactgca	acctccgcct	cctgggttca	agtgattctc
	gtgcttcagc	ctoccaaagta	cctgggacta	caggtgtgca	ccaccacacc	aggctaattt
55	tttgtgtatt	tttactagag	acagggtttc	accatgtttg	ccagcatggg	cttgaacgcc
	tgacctcaga	tgatccaccc	accttggcct	cccaaagtgc	tgggattaca	ggcatgagcc
	accacggcca	gcccacaatg	atattacaaa	cctatttaaaa	atgatactta	gacagaattg
	tcagtattat	tcaagaacat	ttaggctata	ggatgttaaa	tgacaaaagg	aaggacaaaa
	atataatatgt	atgtgacct	accataaaaa	aatgaaatat	tcacagaatc	agatctgaaa

```

acacatgtcc cagactgcat actgggggtcg tcatgagggtg tctccttctt tctgtgtact
tttccttgaa tgtgcacttt tataacatga aaaataaagg tggggaaaaa agtctgaaga
tctaagattg gagagagggtg accttttcagg aaggagagact agaaagaaat atgtgcctgg
5 ttttagagccc tgggtcctgcc ggccctgttc cagggcataat ttccatttcc cagatctcag
tttttctgt ctgtaaaatg ggagagagag gaaaggatgg agagagggaag aaggaaggga
ggagggagga gagaacaggc caacttcatac agcgtgggaa ggggtgtgaa agtgtttctg
agcatctcac gagtgacaag tgaggaggga ggctggcggg tttcagaggg attgggatga
cagtagacag gacacagggg tcccacgggg gtctgccaga agtaagcaaa cagtgcogga
10 ggaagatggg ggcacctgct ccccaagaag ggagggaag gaacctcggg aagcgggtag
gatgaggggg gagtccctctg tgactcagag cctggccaca gcccagcca tctaactca
aagatccctct gtgtgggcac acctcagacg ctgctgaccg aggagccact ccagcccagg
acacgccctc ctacctgttc ttctctgttt tctcccagaa ttccctcccc accaagatcc
tccagatcct tccccctctt atctcatctc cctctgagtc tctcctaacc caggcaccac
15 agccctgtca tattgcagaa attctgcagc cgctaattct gattctoccca tataggaggc
taacacagaa aacgcaggag tccaggcccc cagcccctcc ttccctcagac ccaggagtcc
agaccccccg ccccaacccc tcctccctca gaccaggag cccagggtccc cagccccctc
tgtttctggg cctgtcaagt ttaagaatgt caaacatttt cgaccagtca ttcccctgaa
20 gtttttagcaa cattttctct ctcttctgca aggcactcca acattcaatc tgggaatttta
aaaagtaaca aaacattgca ttgtcactaa gtcagcctgg agatccctgg ccctggccct
ctgctctcct atacgcaagc tacaggtaga ttggtttgca atgactgaga tgggtactaat
gttgattttt tttaagtaat tcatttttct ttgggtaagc agtatagtgt ggtagttaag
ggactagctc tggatcttgg cttcttgggt tcaaatccca gttctagtcc ctacaagcta
25 ttttctttaa agctcattac ttcccctgtc cctgttcttt catccttgaa atgggagaaa
aagcacctac tttctagggg tattacagag attcaataag ttaatatata gaaagtgtc
aaacattgt

```

SEQ ID NO. 13

Hk8 amino acid

```

30 MGRPRPRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHECQPHSQ
PWQAALFQGQQLLCGGVLVGGNWWLTAAHCKKPKYTVRLGDHSLQNKDGPQEIPVVQ
SIHPHCYNSSDVEDHNHDLMLLQLRDLQASLGSKVKPISLADHCTQPGQKCTVSGWGTV
TSPRENF PDTLNCAEVKIFPQKKCEDAYFGQITDGMVCAGSSKGADTCQGDSSGGPLVC
35 DGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKIIGSKG

```

SEQ ID NO. 14

KLK8 nucleic acid

40 CDS 35..817

```

gtgaccccg ccttggtatc tgggaagacct caccatggga cgcccccgac ctctgtcggc
caagacgtgg atgttcctgc tcttgctggg gggagcctgg gcaggacact ccagggcaca
45 ggaggacaag gtgctggggg gtcattgagt ccaaccccat tcgcagcctt ggcaggcggc
cttggtccag ggccagcaac tactctgttg cgggtgtcct gtaggtggca actgggtcct
tacagctgcc cactgtaaaa aaccgaaata cacagtacgc ctgggagacc acagcctaca
gaataaagat ggcccagagc aagaaatacc tgtggttcag tccatccac acccctgcta
caacagcagc gatgtggagg accacaacca tgatctgatg cttcttcaac tgcgtgacca
50 ggcattccctg ggggtccaaag tgaagcccat cagcctggca gatcattgca ccagcctgg
ccagaagtgc accgtctcag gctggggcac tgtcaccagt ccccagaga attttcctga
cactctcaac tgtgcagaag taaaaatctt tcccagaag aagtgtgagg atgcttacc
ggggcagatc acagatggca tgggtctgtg aggcagcagc aaaggggctg acacgtgcca
gggcgattct ggaggcccc tgggtgtgtg tgggtgcact cagggcatca catcctgggg
65 ctgagacccc tgtgggagg cgcacaacac tggcgtctat accaacatc gccgctacct
ggactggatc aagaagatca taggcagcaa gggctgattc taggataagc actagatctc
ccttaataaa ctcacaactc tctggttc

```

SEQ ID NO. 15

KLK8 nucleic acid

5 CDS join (<1..39, 418..712, 878..>946)
 Exon <1..39
 Exon 418..712
 Exon 878..946

10 tcttcggttc ccggttactg gcagcagccc cctcctccca caaaagatca ggttccaagc
 ttctcctttt aaaagtactt agaattcagc cccagctct ctctccctc acaccagga
 atccaggccc ctagccccc ctcctcaga cccaggagtc ctggccccta gcagcccct
 cctccctcag acccaggagt ctgggcccc agccctcct cggtcagacc taaatcccag
 gtcccagtc ctcctccctt agatttagga gtccaggccc ccagcctct ctcctcaga
 15 cccaggaatc caggccccca gcctcctccc ctctcagaac taaaatcttg gccccagcc
 ctttatgttt cagatcgttag agtctcagca ccgagtcct cctctcccta gcctcaggag
 tctgagattc cagccccctc tccctcaaga ttccacgttc aatccccctc gccccttctc
 actcacaccc agtggtccag ttcccagaag ctccccaggc tctagtgcag gaggagaagg
 aggaggagca ggaggtggag attcccagtt aaaaggctcc agaatcgtgt accaggcaga
 20 gaactgaagt actggggcct cctccactgg gtccgaatca gtaggtgacc ccgcccctgg
 attctggaag gtgaggtgca gaggtactca gatacagaca tcaggccccg gaccctcctt
 ctccagattc caggacccca gcctcagatg cccttctctg tcgagatcca gcagcttgga
 ccccggttcc ctcctctccc taatttagga gtcccagctc ccagctccct gtccctcag
 acccagacat cgaggactcc cccctccctt ggaatgtagg aatccagtc cccagcctcc
 25 tccttctctc agagaagccc agaacagccc cagatactct cggctgcctc cccagtgcct
 aaatccagaa ctgggagctc aggtcctctc ttctgttta ccggccccgc cctctccatt
 tcccagacct caccatggga cgccccgac ctctgtcggc caagacgtgg atgttctgtc
 tcttgctggg gggagcctgg gcaggtgagg aggttgctgg aggcctccgg aggggaggga
 tctgaaggca gcagtggcgc tggggagtct gtgggaatgc cgcgggggtt atgtgggtgc
 30 gtgtgcacgg atgtgaagag tgcgatacgg tgcaggagcc tctgtgggtt ttctcaggg
 tggacagagg caagaaacag gtagcagcag gtaggagtag gtccgtgat gctgtaaat
 gtctgaatag ctacagcctt tgggggctgc ttgcttggg gcatagattc acctgggagt
 actcggggcc tgtagactca tgtggaagca tgtgggggca ttcttgggtg tgtgactctt
 gtagatgac acatggactg aaatgagtgt cccgtgtgg cagcgtgtgg aagcctggac
 35 ctctcacta agttgtatgc ggagaacttg ccgtgtgtcc atttgaacc acagtggcct
 tcccagccct cgcactgccc cagaggggtg cgatccaacc ctctccctcc tgotgcagga
 cactccaggg cacaggagga caaggtgctg ggggtcatg agtgccaacc ccattcgag
 ccttggcagg cggccttgtt ccagggccag caactactct gtggcgggtg ccttgtagg
 ggcaactggg tccttacagc tgcccactgt aaaaaaccgt gagtggatga tgggggcaga
 40 ggtcagctgg ggcttaagga aagagggggc tggggtttcg actcaggaag gagagagctg
 aggactggac toctgggtct gaaggaggag ggggtggg gcaatacccc tgctgggtc
 ccaaactatc cccaccatta caggaaatac acagtacgcc tgggagacca cagocacag
 aataaagatg gccagagca agaaatacct gtggttcagt ccattcccaca cccctgtac
 aacagcagcg atgtggagga ccacaacat gatctgatgc ttcttcaact gogtgaccag
 45 gcatccctgg ggtccaaagt gaagcccatc agcctggcag atcattgcac ccagcctggc
 cagaagtgca ccgtctcagg ctggggcact gtcaccagtc cccgaggtag tgggcttgtc
 cactaatggg agggagagga ggagctggtt ggccagtg aaccaagct attggcaaag
 cttggtcccc cagagggaga caaagaagg aaagtgatca tgatgttgag attcacaag
 aggagtcata tgagaagcct cgaagatctg actactaaca agagtgtgga gagaaagaac
 50 caactagcag attgttaagc agggccagaa aagcccatcc tgtatggcgg agagcactac
 cctatgggac ctttgggtgt ataggatttt atataatctt aacacacttg ggatattttg
 gacttctcag agggccagga aaacaggctc ctaagcacct tctccccac ctcctcctt
 tttttttttt tttttgaggg ggacacagt ttactcaatc tccactcaatc gcccaggtg
 gagtgcagtg gcacaatctc agctcactac aatctctgcc tctgggttt aagcattct
 55 cgaacctcag cctcccgagt agctgggatt acaggcacc gccaccacgc gagctaattt
 ttgtatttgt ttagagaca ggggttcgcc atgttgacca ggctgggtct gaactcctga
 cctcaagtga tctgcccgc tcggcctacc aaagtgtg gattacaagc atgagccacc
 gcccctggcc ttttctctc tcttgaacc aggaaacct ggccctggc acctcttccc

	tcagaccccg	gagtctaggc	ccttcttctc	ccaggaccca	gcagtcctag	ctccctcttg
	actctggacc	ccaaaatctg	gacctccaat	gaagctgtcc	ctttgggact	ccagaatcca
	gagagcccct	cgccttcctt	cacagtgaag	acattgggac	tcacctaaga	gagtcgaagg
	gcttctccag	gaagtggcaa	agtcagcatt	caggtccctg	cctgcctcac	tcctgctctg
5	aatgctttgg	atgagacagt	ttggcgctgt	ggaaacacac	gtgctcgcaa	atcaagtaga
	tcagttcaaa	ctatggctct	gccctttctc	cctgggcaaa	ttcctttcca	tctctgagcc
	tcacttttct	tatctgcaaa	atgggaatca	ttaaccaata	gatttttttag	ctacgtgagg
	gttcagggtta	catcagtttt	cttcattgtg	ttctcctagt	gctcagaatg	gtgcctgcta
10	catggtagat	gctcaataaa	tatttggtga	atggatgacc	tgatgaataa	ataatgaaat
	gaatgaatgc	atctctcctt	caaagcgctg	ttgtgaggat	taaatgagat	tatgagcata
	ttgcttttgg	cacgtagaga	tgccagatgg	aagtttttcc	ccctcaaaga	ggcttttgag
	aagtctattc	ctcaaaagag	gttaataaaa	aagatcaatt	ccaccttcaa	tcattaattc
	aactcttatt	tactgagcac	ctagtatgoc	tgaggtgctg	ttgcaggcgc	tggatataca
	gccatgagca	aactgtacaa	agtccttgtc	gttatggagt	tgcaagctag	gtgggagaaa
15	tagacaataa	acaaatacac	ataaaataaa	acattaggca	aagtgctgta	aggaaatatt
	gtggagcgta	aaaggatagg	gagtaatgga	gggatgggat	tttttaattc	gagtggtcag
	ggaaggcttc	caggaggagg	tgacacttgg	aaggagcagg	atttagcaca	gattgaggac
	ccggttctct	gagggcagga	agactggaaa	gagagggtca	gaggtatcat	aaaggggcac
	ctccaagtaa	ccccagccc	cttgatttga	aaatggctca	gggtaagaaa	aaacacgtga
20	gatccaaggg	cccctctcta	gatggagaaa	gccaatagc	aacaagtaca	gcttcgttta
	atgtggtgag	aagtgatgtc	ccctgtgcac	agtgtcagaa	aattccccc	tgagctgga
	aaactcccc	attacatcct	ggaaagaaa	ggggttagat	ccagacaggg	atggaggcaa
	agggtgctc	tctcagggaa	ccttacaacc	tcttccccct	cagagaattt	tcctgacact
	ctcaactgtg	cagaagttaa	aatctttccc	cagaagaagt	gtgaggatgc	ttaccogggg
25	cagatcacag	atggcatggg	ctgtgcaggc	agcagcaaa	gggctgacac	gtgccagggtg
	agcaatttct	gaaatccttc	tcctcacaca	tcctctattg	ccctctcgag	gttcaaggct
	tggatggggg	tgggggtggg	aagggagtga	cccaaagac	ttggcaccgg	gagtggtcac
	ccctatctct	acggattggg	agccagggtc	agagaagcca	aactctctct	ctgaaagtca
	cactgcatag	ggattaggag	cattgaattg	ctgctgctgt	tttctttcga	acgcttaact
30	acagacttgg	atgcagagtg	gggggctata	gaggggtggg	tagatgtcca	gcaaggagag
	agtctgaggg	tagagttagg	acgagttgcc	tgctctctgg	ttcccagcat	tacatcgtag
	aatttgtgcc	gactttacca	ccggcccggt	ggtcagccac	ttttctttga	gaccgggcta
	gaatcccaag	gctggttccc	ttctcctgtg	attggttcc	tgaggagaca	tggtgtcttc
	ccagttggct	ggagtaaatg	gtgccattga	cttcagtgct	tcctcaatga	gtccaatcc
35	ctttctcttg	ggatttgtca	tagatataac	tctctcttct	aactgcaact	tcgttgttcc
	tagcactttc	ccctggcttt	gtccacttcc	tggaagcccc	accacctttg	ccatgactg
	gtccctaaat	acaatgcttt	cttccccatt	ggccaaaaat	ggagtctgtt	ccatcagcga
	tactgccatg	aaagccagtc	tctggattgt	tctgtagaga	tagtggtctc	ttccacaaat
	atttcagcca	tggtctcttg	gggatattgg	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt
40	gtgtgtgtgt	gtgtccataa	gaatcttgat	cctttctcct	atgtttggca	actgcaatca
	atgatgcctt	cactattggc	caggaacaga	ggaaacttca	gctcagtcct	ctccaagtaa
	tcctactgtg	cttctccctg	gattggaccc	tcgagaactc	tttttttttt	tttttttgaga
	cagggtctcg	ctctgtcccc	caagctggaa	tgacgtggca	caatcttggc	tcactgcagc
	ctctgectcc	cagttcaagc	aattctccca	cctcagcctc	ccgagtagct	gtgattacag
45	gtgtgcacca	ccacaccag	ctagtttttg	tatttttagt	agagacaggg	attcaccatg
	ttggccaggc	tggtctggaa	cgcctgactt	caagtgatct	accgcctcgg	cctcccaaag
	tgccgggatt	gcaggtgtga	gccaccaagc	ctgtctggga	tttcattctt	tccccctctc
	tgctagtggt	ttgaccacta	cccttagaca	ccatgtctgt	ctgtacatgg	aagccccaa
	ccctgtcctg	actggtctca	ggggacaaat	cttttaccct	cattggctac	aggggaccaa
50	tcagtccaaa	gaactggtaa	aacgtctggg	cagcaggaaa	agggacgttg	tggacatctc
	agatgcaagg	ctgttctcat	tctccctgtc	tagggcgatt	ctggaggccc	cctgggtgtg
	gatggtgcac	tccagggcat	cacatcctgg	ggctcagacc	cctgtggggg	gtccgacaaa
	cctggcgctc	ataccaacat	ctgcccgtac	ctggactgga	tcaagaagat	cataggcagc
	aagggtgat	tctaggataa	gcactagatc	tcccttaata	aactcacaac	tctctgggtc
55	cttgccctgtc	tctgttttgg	ccctgtgggg	agggctggat	ggggatccgg	gattgttcc
	gctggcagac	taggtgggga	tgtgcagaaa	ccaagttctc	atgggtcact	tggccatcac
	cactgcctaa	agtgatccct	cgttttctgg	aagaacttgg	gtaagagctt	tatttcaggg
	gagaaaatac	atacaaaagg	cttcaaacat	tcagtgggct	ggatgtgtaa	agacagtttt
	gaaagagttt	gtgttttagt	ttcctgagca	aagcattttac	aagctttgga	gataaaattt

tcccccttta aaaaataatt

SEQ ID NO. 16

5 Hk10 amino acid

MRAPHLHLASAASGARALAKLLPLLMAQLWAAEAALLPQNDTRL
PEAYGAPCARGSQPWQVSLFNGLSFHCAGVLVDQSWVLTAHCGNKPLWARVGDDHLL
LLQGEQLRRTRTSVHVKYHQSGPILPRRTDEHDLMLLKLARPVVPGRVRLQLPY
10 RCAQPGDQCQVAGWGTAAARRVKYNKGLTCSSITILSPKECEVFYPGVVTNNMICAGL
DRGQDPCQSDSGGPLVCDLTLQGILSWGVPYPCGSAQHPAVYTQICKYMSWINKVIRSN

SEQ ID NO. 17

15 **KLK10 nucleic acid**

Gene 1...1580
CDS 220...1050

20 catcctgccacc cccctagcct tgctgggggac gtgaaccctc tccccgcgcc tgggaagcct
tcttggcacc gggacccgga gaatccccac ggaagccagt tccaaaagg atgaaaagg
ggcgtttcgg gcaactggag aagcctgtat tccagggcc ctcccagagc aggaatctgg
gacccaggag tgccagcctc acccagcagc atcctggcca tgagagctcc gcacctccac
ctctccgccg cctctggcgc cgggctctg gcgaagctgc tgccgctgct gatggcgcaa
25 ctctggggccg cagaggcggc gctgctcccc caaacgcaca cgcgcttga cccgaagcc
tatggctccc cgtgcgcgcg cggtcgcag ccctggcagg tctcgcctct caacggcctc
tcgttccact gcgcgggtgt cctgggtggac cagagttggg tgctgacggc cgcgactgc
ggaaacaagc cactgtggc tcgagtaggg gatgaccacc tgctgcttct tcaggggagag
cagctccgcc ggaccactcg ctctgttgtc catcccaagt accaccagg ctcaggcccc
30 atcctgcaa ggcaacgga tgagcacgat ctcatgttgc tgaagctggc caggcccgta
gtgctggggc ccgcgctccg ggccctgcag ctccctacc gctgtgctca gcccgagac
cagtgccagg ttgctggctg gggcaccacg gccgccgga gagtgaagta caacaagggc
ctgacctgct ccagcatcac tatcctgagc cctaaagagt gtgaggtctt ctaccctggc
gtggtcacca acaacatgat atgtgtgga ctggaccggg gccaggacc ttgccagagt
35 gactctggag gccccctggt ctgtgacgag accctccaag gcatcctctc gtgggggtgtt
taccctgtg gctctgcca gcatccagct gtctacacc agatctgcaa atacatgtcc
tggtatcaata aagtcatacg ctccaactga tccagatgct acgctccagc tgatccagat
gttctgctcc tgctgatcca gatgccaga ggctccatcg tccatcctct tctccccag
tcggctgaac tctccccctg tctgactgt tcaaacctct gccgccctcc acacctctaa
40 acatctcccc tctcacctca tccccccatt tatccccatt ctctgcctgt actgaagctg
aaatgcagga agtgggtggc aagggttatt ccagagaagc cagggaagcc gtcacaccc
agcctctgag agcagttact ggggtcacc aacctgact cctctgccac tccctgctgt
gtgacttttg gcaagccaag tgcctctct gaacctcagt ttcctcatct gcaaatggg
aacaatgacg tgcctacct ttagacatgt tgtgaggaga ctatgatata acatgtgtat
45 gtaaattctt atggtgattg tcatgtaagg cttaacacag tgggtggtga gttctgacta
aaggttacct gttgtcgtga

SEQ ID NO. 18

KLK10 nucleic acid

50

Gene 1..5574
mRNA join(48..120,605..701,2455..2635,3589..3863,4195..4328, 4793..5474)
CDS join(614..701,2455..2635,3589..3863,4195..4328,4793..4945)
Promoter 1...47
55 5'UTR join(48..120,605..613)
exon 48...120

exon 605...701

	ttgggggtcaa	aagggaaggt	cccgccaggg	gtccctgggc	agaggatacc	agcggcagac
	cacaggcagg	gcagaggcac	gtctgggtcc	cctccctcct	tcctatcggc	gactcccagg
5	tgaagctacc	tgaccccac	ccgggttggg	gtggattgcg	agaggatggg	tgggaacccc
	cgggccacag	gcaggagccg	gcttagagcc	tcggttttctc	cactgcggga	cgcggaagtc
	ccccgcgtgt	gagggttgaga	aagaggctcc	cactggctcc	gagcctcgga	tccccacccc
	gcgcgtgaag	gagggggaaa	cctcggggcg	ggcgtggctg	cagccagggt	agctggggcg
	cggagagcgc	tccactcggg	cacagggagg	acgggaagat	gccgcgaggg	gcgtcattag
10	ggtaattgtg	cccattaccg	tgttccagcc	cgaatctccc	gtctgccagc	ccctgggggt
	atcggctgca	ggtcaaaagg	gtgcttgogg	gctgggggtg	cccaggctgg	gcggatgagc
	gcggcgagggt	gggggtctct	aacggagcat	ctgttttaac	ccgcccctgc	acacacccca
	gcagatcctg	gccatgagag	ctccgcacct	ccacctctcc	gccgcctctg	gcgcccgggc
	tctggcgaag	ctgctgccgc	tgtctgatgg	gcaactctgg	ggtaagggtg	gggacagggg
15	gcggggagag	gcgcgggtgg	gaggcacogg	cgggagggca	atgtgttccc	gtcaccaagc
	cccgcgacac	tctcctcccc	cgagacccca	gcacccaccc	agcgtctccg	agcacggccc
	gcccccaagt	cagctggggc	cttcttctgg	ctcgccccc	gggtgaccgg	cccactcag
	gcccgtgtccg	atttctgcca	cccggtatcc	cgctcttccg	tggactcttc	ggcgtgttct
	ttctcctcct	cttacgccgc	cccatccogg	ccccgctcca	ttttacacta	ccgtgttttg
20	ttttgtttgt	ttgtttgttt	gagacggagt	ctcgctctgt	cgcccaggct	ggagtgcagt
	ggcgcgatct	cggctcactg	caagctccgc	ctcccgggtt	cacaccattc	tcctgcctca
	gcctcccagag	tagctgggac	tacaggcgcc	cgccaccaag	cccggctagt	tttttgtatt
	tttagtagag	atgggatttc	accgtggctc	cgatctcctg	gcctcgtgat	ccgcctgcct
	cggcctccca	aagtgcctgg	attacaagcg	tgagccaccg	ctcccggcca	caccacagtg
25	ttttatcctg	agtcttgcc	taccgctttt	tgcctctctc	cctcactttt	tttcttctc
	tttccctttc	tctctctttt	ttctttttct	ttctttcttt	tctttctttc	gttcttcttt
	tccttctctc	ctttccttcc	ttcttttctg	ttgtttcttc	tttcttgttt	tactttctct
	cttatttttt	cttctttctt	tcttgttttt	tttctttctc	tctctttttc	tttcttctct
	tttttttttt	tttttttgag	acagggcagt	gctctgtctc	cgtggctgga	gtacagtggc
30	ccaatcagag	ctcactgcag	cctcgacctc	ctgggctcaa	gcgatactca	gcctccagag
	tagctggtac	cacaggcatg	caccaccaca	tcgggctttt	tttttttttt	tttttttttt
	tttttttgag	acagggtctc	actctgtcgc	ccagactgga	gtgcagtggc	ccaatctcgg
	ctcattgcaa	cctccacctc	ctgggctcaa	gcgatcctcc	cacctcagcc	tcccaggtag
	ctgtgactac	aggcgcatgc	ctccgcgaat	acttttttgt	tgttgttgtt	gtttgtttgt
35	ttttgtagag	actgggtctt	gctgtgttgc	ccgggctggg	cttgaattcc	cgagctcaag
	cgttccaacc	gcctcgggtt	cccaaagtgc	tgggattaca	ggcgtaaaac	actgcgcccc
	accctctctc	tggttttcaa	tcccgttttg	ttattcacac	cccttctctc	ccccgatccc
	cgagttctat	ccccgcaccc	ttacctcccc	gccgcgttca	atccccgccc	ctctatcgac
	cagcgacgtt	ctagccagct	ctccaggcgc	gctgcgttca	gtccctgccc	tccagaccca
40	ccctattctg	tctcattact	ccacgctacc	ctatcccagc	ttccttccac	tttcacgcgc
	ttcttctcct	cccattcctt	cgggtgcacg	gaaaccccca	atatttccct	accaccctcg
	cgttctggct	gcgtcccccg	tcccgaacgc	cagtccagtg	ccacagccca	gcctccaacc
	caaaccacga	ccctgccttc	cgtgctttga	ttccgtcccc	tttctttctc	ccagccgcag
	aggcggcgct	gctcccccaa	aacgacacgc	gcttggaccc	cgaagcctat	ggcgccccgt
45	gcgcgcgcgg	ctcgagccc	tggcagggtc	cgctcttcaa	cggcctctcg	ttccactgcg
	cgggtgtcct	ggtggaccag	agttgggtgc	tgacggccgc	gcactgcgga	aacaagtagg
	aggagatcca	tccccgagga	cgccacgggg	ggctgtggag	gcggcctcca	gggaggagcg
	ggcagggcgg	ggtctctcgg	aactcccaca	gctggagggtg	cggccccggg	tgtcttcca
	gcaggaggag	agaccgggct	tgcgctgtgg	ccgccagggg	acggtgtggg	ccttttccgc
50	atctctgggc	ccgtcactcc	tctcccgcgt	attctccttg	agctctagga	ggaggtgggt
	gcctgcaacg	gatagaagcc	agggtoogg	gtggtcagga	atttagaact	aaaggaaaag
	gttcacttctg	tgagtccccg	ttgaaggagg	aagggttggg	tattaccaca	gagaaaatgt
	ggagttgggc	tgggctcggg	ggctcacgcc	tgtaatccca	gcactttggg	aggccaaggg
	gggtggatca	cctgaggtca	ggagttcaag	accagcctgg	ccaacatggt	gaaaccactt
55	ctcaactaaa	aataaaaata	aataaaaaat	aaaaaatttt	aaaaaaaatta	aaaaaaaata
	aaggctgggt	gtggtggcgg	gtgcctgtaa	tcccagctac	tcaggaagga	ggctgaggca
	ggagaactgc	ttgaacctgg	gaggtggagg	ttgcagttag	ccgagattgc	gccattgcat
	tctagcctgg	gtgacaagag	cgaaactctg	tttcaacaaa	gaaaagaaaa	agagaaaaga
	aaatgtggaa	ggcttaccta	ggtgtccagg	ccccagccc	tcctcaattc	ctgcagatcc

```

tcagagctca aacaactgat tctctctccc catgtccact gaggtccct tctcccacaa
ggccctcttc cctcagactc ttcctatctc caggccctgc ttcactgccc acctgctttc
ccagtccttg tgaaggggtt gccttcacat gcctcttcct tcccccaggc cactgtgggc
tcgagtaggg gatgaccacc tgctgttctc tcagggcgag cagctccgcc ggacgactcg
5 ctctgttgat catcccaagt accaccaggg ctcaggcccc atcctgcaa ggccaacgga
tgagcacgat ctcatgttgc taaagctggc caggccccgta gtgccggggc cccgogtccg
ggccctgcag cttccctacc gctgtgctca gcccgagag cagtgccagg ttgctggctg
gggcaccacg gccgcccga gaggcaagag ctggggctct gaggccagaa cctcaggagg
agggggctga gggcctgaac ccctgggtct gaggaaggat gggctgggga ctggattcct
10 ggatctgagg gaggacgggc tggggctcta gatgcctggg tctgtgagtc tgaggggagg
aggggctggg ggctggact cctgggtcta agtggggagg ggctggggcc aggatcttg
agtctgaagg aggaggggct ggggcttagg atagaaacgg tctgtatct ggactcctgg
ctccccaagg attgggggct ggaccaggg attactggca tattctccct tcagtgaagt
acaacaaggg cctgacctgc tccagcatca ctatcctgag ccctaaagag tgtgaggtct
15 tctaccctgg cgtggtcacc aacaacatga tatgtgctgg actggaccgg ggccaggacc
cttgccaggat agggctctgaa cagggagagt ctctgactcc tgggagggag gacagggagg
ttatgggaaa agagcagacc ctgtgcccga tcccaaactc cattcccaa cccatccttg
acccaactc ttaccagac ctaacccct cctcatccct atcctcaatc ccatttccat
cctaaccoca ccccatccct atctccaagc ccattttcat cccctcacct tccatgaact
20 acaatcccaa cccaagtctc actgtgcctt cattctcatc cccagccca acctccata
acctgaagtc cacctccatt cctaccttcc agctcatacc taattccaac cccatcccat
cctcgtcttt atcccaaccc aaccccttcc ttccccacca ctgccccaga tcccaaagtg
acagctctca cgttggcaca tttatttgat ctctccttcc tgccaccccc agagtgactc
tgaggcccc ctggtctgtg acgagaccct ccaaggcatc ctctcgtggg gtgttttacc
25 ctgtggctct gccagcatc cagctgtcta caccagatc tgcaaataca tgtcctggat
caataaagtc atacgtcca actgatccag atgctacgct ccagctgata cagatgttat
gctcctgctg atccagatgc ccagaggctc catcgtccat cctcttctc cccagtcggc
tgaactctcc ccttgtctgc actgttcaaa cctctgccgc cctccacacc tctaaacatc
tcccctctca cctcattccc ccacctatcc ccattctctg cctgtactga agctgaaatg
30 caggaaagtg tggcaaagg tttattccaga gaagccagga agccggtcat caccagcct
ctgagagcag ttactggggg caccacaact gacttccctc gccactcccc gctgtgtgac
tttgggcaag ccaagtgcc tctctgaacc tcagtttct catctgcaa atgggaacaa
tgacgtgcct acctcttaga catgttgtga ggagactatg atataacatg tgtatgtaaa
tcttcatgtg attgtcatgt aaggcttaac acagtgggtg gtgagttctg actaaagggt
35 acctgttgat gtgatctgac cagctcccg tgaaagcgtg tgtccaggga agaagtgcac
agggtagccc ccaatcccaa ccttccatcc ccaaccctta gggatgatgg aaga

```

SEQ ID NO. 19

40 Hk11 amino acid

```

MQRLRWLRDWKSSGRGLTAAKEPGARSSPLQAMRILQLILLALA
TGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAHCLKPRYIV
HLGQHNQLKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSIWAVRP
45 LTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRANITIIHQKCNAYPGNITDTM
VCASVQEGGKDSCQGDSSGGLVCNQSLOGIISWGQDPCAITRKPGVYTKVCKYVDWIIQ ETMKNN

```

SEQ ID NO. 20

50 Hk11 amino acid

```

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTR
LLCGATLIAPRWLLTAAHCLKPRYIVHLGQHNQLKEEGCEQTRTATESFPHPGFNNSL
PNKDHRNDIMLVKMASPVSIWAVRPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHT
LRANITIIHQKCNAYPGNITDTMVCASVQEGGKDSCQGDSSGGLVCNQSLOGIIS
55 WGQDPCAITRKPGVYTKVCKYVDWIIQETMKNN

```

SEQ ID NO. 21

KLK11 nucleic acid

```

5  aggaatctgc gctcgggttc cgcagatgca gaggttgagg tggctgcggg actggaagtc 61
   atcggggcaga ggtctcacag cagccaagga acctggggcc cgctcctccc ccctccaggc 121
   catgaggatt ctgcagttaa tcctgcttgc tctggcaaca gggctttagg ggggagagac 181
   caggatcatc aaggggttcg agtgcaagcc tcaactccag ccctggcagg cagccctgtt 241
   cgagaagacg cggctactct gtggggcgac gctcatcgcc ccagatggc tcctgacagc 301
   agcccactgc ctcaagcccc gctacatagt tcacctgggg cagcacaacc tccagaagga 361
   ggagggctgt gagcagaccc ggacagccac tgagtccttc cccacccccg gcttcaacaa 421
10  cagcctcccc aacaaagacc accgcaatga catcatgctg gtgaagatgg catcgccagt 481
   ctccatcacc tgggctgtgc gacccctcac cctctcctca cgctgtgtca ctgctggcac 541
   cagctgcctc atttccggct ggggcagcac gtccagcccc cagttacgcc tgcctcacac 601
   cttgcgatgc gccaacatca ccattcattga gcaccagaag tgtgagaacg cctacccccg 661
   caacatcaca gacaccatgg tgtgtgccag cgtgcaggaa gggggcaagg actcctgcca 721
15  ggggtgactcc gggggccctc tgggtctgtaa ccagtctctt caaggcatta tctcctgggg 781
   ccaggatccg tgtcgatca ccgaaagcc tgggtctctac acgaaagtct gcaaatatgt 841
   ggactggatc caggagacga tgaagaacaa ttagactgga cccaccacc acagcccatc 901
   accctccatt tccacttggg gtttgggttc tgttactct gttaataaga aaccctaagc 961
   caagaccctc tgcgaacatt ctttgggcct cctggactac aggagatgct gtcacttaat 1021
20  aatcaacctg ggttctgaaa tcagtggagc ctggattcaa attctgcctt gaaatatatt 1081
   gactctggga atgacaacac ctggtttgtt ctctgttgta tccccagccc caaagacagc 1141
   tcctggccat atatcaaggt ttcaataaat atttgctaaa tgagtg

```

SEQ ID NO. 22

25 KLK11 nucleic acid

gene 2313..7622

30 mRNA join(2313..2398,4189..4263,5061..5217,5545..5810,
6627..6763,7158..7622)

CDS join(4224..4263,5061..5217,5545..5810,6627..6763, 7158..7310)

```

35  tgataaatagt gttctctctc ctcatgtgtc agggccccag ccattgtcct tgagagaatg 61
   ctcgactctt tatgttgtct tgacagcctc ccctgagatt ggtcattaat gactgtgtct 121
   tctctcctca ttgggtcaggg ccccagccat tgtccttgag agaacctctg tcctttatgg 181
   agttccaccc ttcttccctg ggattggccc ctagagacag tggttcttct cttttgggta 241
   gccattgcca ttgtcctccg ggaaagtgtat tatactcttt tgtctaataa ccagacttgg 301
40  agccctcccc aaggcccagg actgggttga agggttgggg aggaaaacag aaataagatg 361
   tctcccttgt tcagacagta cttctcttcc cttccagggt gattctgggg gccccctggg 421
   gtgtggggga gtccttcaag gtctggtgtc ctgggggtct gtggggccct gtggacaaga 481
   tggcatccct ggagttctaca cctatatatt caagtatgtg gactggatcc ggatgatcat 541
   gaggaacaac tgacctgttt cctccacctc cacccccacc ccttaacttg ggtacccctc 601
45  tggccctcag agcaccaata tctcctccat cacttccct agctccactc ttgttggcct 661
   gggaacttct tggaaacttta actcctgcc gcccctttaa gaccacgag cggggtgaga 721
   gaagtgtgca atagtctgga ataaatataa atgaaggagg ggccatgtct gtccatttga 781
   agtctcatg ctggttgaga ctggaagaag gactcagcag tttccctatc tcataggagt 841
   agaaacagag ctcaaataag gccaggcaca gtggctcaca cctgtaatcc catcactttg 901
50  ggaagctgag gcaggtggat cacctgaggt caggaactcg ggaccagcct ggtcaacata 961
   gtgaaacccc aactctacta aaaatgcaaa aattagccag gcatggtggc gcatgcctgt 1021
   aatcccagct actcaggagg ctgagacagg agaatagcat gaaccctga ggcagggt 1081
   gcagcgagcc gagattgaac cattacactc cagcctgggc gacagagcga gactccatct 1141
   caaaaaacaaa caaacaacaaa acccagtgct caaataggat gaggtcttcc cctgagtagt 1201
55  tactcagaaa tggagtagaa aaagttactt ttaataatat aggccgggtg cagtggcca 1261
   cgctgtaat cccagcactt tgggaggccg aggtgggagg atggcttgag ctgagatttc 1321
   gagatcagcc tggcaacaca gtgaaatctt gtcactacaa aaacacaaaa aattagctgg 1381

```

	gtgtgggtggt	gcgtgcctgt	agtcccagct	acttgggaag	ctgaggtggg	aggatcaccc	1441
	gagccgggga	ggtggaggct	gcaaagagcc	gagatcatgc	cactgcactc	cagcctgggc	1501
	aataaagtga	gaccttgtct	caaaaacaaa	aaccagcaa	tataaataag	acacatgttt	1561
	cttcatctgg	cataatagaa	atagtgccca	gagcttataa	gcttttcaag	agtccacaaa	1621
5	agacccgaaa	aagaaaaaga	aaattgttag	ctccaaaata	ccagatgaaa	gctgcaaaagt	1681
	caacattttat	gaccatttta	tccaatgttc	ataaaacgta	gcattctttc	cactagccaa	1741
	ctgcagtttta	ctttcttgta	atgaagcata	cattgtatct	ttaatgtggg	acgtggcttt	1801
	gttctaataa	gacgaagggt	ggagtgcagg	cttggaaagc	aggagagctc	agcctacgtc	1861
10	tttaatcctc	ctgcccaccc	cttggattct	gtctccactg	ggactcaaga	ggtgaggaga	1921
	gaccatctcc	ccaaatgcac	tgaagggaaa	ctggaggagg	gagggagtga	gggtgatca	1981
	taccagcgga	ggcacatttg	ctgagccccc	ccgcagtctg	ctctttccaa	gtggaccctc	2041
	ctggaagcct	gatcccaacc	tcccctgcaa	gcaggtctgt	cacccccatc	tctcagatga	2101
	agaaactgag	ccttgcaagg	gtggagtccc	ttgtcccccac	gtcataaggg	tagtcatagt	2161
	agtaggaaga	ggaagcacct	aggtttgagg	ccagggtctg	ctgctgtcag	aacctaggcc	2221
15	ctcccctgcc	ttgctccaca	cctggtcagg	ggagagaggg	gaggaaagcc	aagggaagg	2281
	acctaactga	aaacaaacaa	gctgggagaa	gcaggaatct	gcgctcgggt	tccgcagatg	2341
	cagaggttga	ggtggctgcg	ggactggaag	tcctcgggca	gaggtctcac	agcagccagt	2401
	aagtgaacag	ctggactcgg	gctgcctggg	cggcaggagg	aagcgggag	gggaagggtc	2461
	agcagaggag	cgaggcccca	gaggagccct	ggggtggagc	acagccaagg	gctctgttcc	2521
20	ctttcctgga	ctcggcttcc	acaggccctg	acctgcctcc	cccaccctcc	ggtcctgccc	2581
	ctgtgcctgg	cagcagcccc	acctgtgtga	catcccagca	cacccccctc	ctccttgcaa	2641
	aggagaagg	agcggcctag	gggaggccag	gggcccacct	gggctggggc	tgtggagagg	2701
	gagtggctgg	gacgggagga	aaaagagaga	cggagattag	atggaagaag	agggatttca	2761
	agacaaattg	ccagagatgc	agtcagagag	actgactgag	agacacaaag	atagaaggaa	2821
25	ttagagaaag	ggccacacag	agccagacag	agagagaaga	gtggagatgg	agacaggagac	2881
	gaggacagag	aaaggcagac	agacacatag	ggacagaaag	agaaaaatca	cacaaagtca	2941
	gaattactga	atgacaggga	atgacacata	gaacgagaca	cagattcaga	gactcagggc	3001
	agggaaagga	aggctgcaga	cagacagaca	gacagaggga	ggctgagaca	cagggagaag	3061
	aggggcttgg	agaggtggca	caggcaggga	gccagtgcct	cagaggcctc	cggggagggc	3121
30	accccgcccc	cccccggccc	ggggcattaa	ggcagggtt	ggaggccagt	catcctgggc	3181
	ccgcccagg	ccgccccct	gccagcccg	ctgcctgggt	cctggcacct	ggcgctccaa	3241
	cccagcctac	ctgctgtagc	tgcgcgcaact	gcgctctccg	ccgccactgg	gccccagag	3301
	ccccagcccc	agagcctgtg	agtccaggag	gaaagggaag	ctgcccctcc	ccgtccagg	3361
	gtcagccctc	cccaaggaca	cctgtcccac	tggggcacc	atttctccct	ctgctctgtc	3421
35	cttctctgct	tgggtggggg	ttcctggcct	ctctctacac	ctctcacctc	cgatggctgt	3481
	ccgcagcctc	agttacctct	aatctccatg	gcttcagctg	ctgagctggc	cctctgctcc	3541
	cacccccgct	ggccagggca	gcgggaggga	ctggccctcc	cctcgaccag	ccccgccag	3601
	ctttgcttgg	ctgtccttca	aaagggcagg	ggtttggcgg	acagggttcc	agcaagccgg	3661
	gtgatggggg	tcccagacat	tgtctggggc	tgagccccct	actcccctcc	agcagacctc	3721
40	aaaggctcca	tatcgctctg	ctgcgaagac	aatgaaaaag	gggtggctac	ggaacggtgt	3781
	ctggttcccc	ttgtccttcc	accccaagct	gctggggcct	ggccagctct	caaggcaaga	3841
	aggaacacac	cctctgacat	gtgcggggga	ggtcccatgg	ctgacttgaa	cagggccgaa	3901
	ccatggcctg	acagctcaaa	gcccctccca	acgacttcca	catggttctt	ggtatctcgg	3961
	aagcttctag	ctgtgaccag	gccctctcca	aggccaccct	agacacctaa	gatataatttt	4021
45	aagtgttttg	agatctgagt	gctgtgagaa	acaggggatt	tccccaacct	tgtttctccc	4081
	aagtggggag	cgggagcagg	tgagggagag	aggagagggc	atgagccagc	ccccccctcc	4141
	cgatttcccc	gtaaaagtga	gcggccccat	gtccctcctt	gttcccagag	gaacctgggg	4201
	cccgctcctc	ccccctccag	gccatgagga	ttctgcagtt	aatcctgctt	gctctggcaa	4261
50	caggatcgca	gggatgggg	gcagggcagg	atcctccctc	ttgaatctct	gggatccctc	4321
	aaccctctgt	gtctggacag	tgacagggct	gattccaaat	tacagaacaa	cccataaggc	4381
	acctgaactg	gagcagtgg	catgaggggc	tggatgccct	tctagataat	ccctttaaat	4441
	gccaaaggag	gagaggtcaa	gggggtcgta	aagggtccc	tggaggggct	gaggaagctg	4501
	gagttggggg	agcagtcact	caaagcgccc	aggacagggg	ctactgacca	accagtatgg	4561
55	aagtataattc	tttttttttt	ttcccagaga	caaagctctt	ctctattgtc	caggctggag	4621
	tgccgtgggt	ccaacacggc	tcactgcagt	cttgacttcc	cgggcttaag	tgatccttaa	4681
	gccatctcag	cttccccgg	agctgggacc	acaggcacct	gccaccaagc	caggctaatt	4741
	gtttaattgt	ttgtagagat	gggggaggag	gtctcactat	gtttgcctag	gctgatctag	4801
	aactcctggg	ctcaagtaat	cctcccacct	tagcctctca	aagtgcctgg	attacaggca	4861
	tgagccactg	catttgacct	tatggaagta	ttttcatcct	ttaatacccg	acccagcat	4921

	ccaggggaac	ccagagggac	accagaccag	ggcccagacc	accactctc	tttctctcct	4981
	ccccaccccc	atttctggga	gtcctcctgg	tctaccacct	ctccttctctg	agcccccttct	5041
	tttgctctca	ccccctccag	ggctttagag	gggagagacc	aggatcatca	aggggttcga	5101
	gtgcaagcct	cactcccagc	cctggcaggc	agccctgttc	gagaagacgc	ggctactctg	5161
5	tggggcgacg	ctcatcgccc	ccagatggct	cctgacagca	gcccactgcc	tcaagccgtg	5221
	ggtgcggggg	ctggggcggt	gccgggggtg	ggggctggga	atggggagat	ggatggagag	5281
	aagctcaggg	ataggggtgc	tggttaaggg	attagagatg	gggatgggta	gtgtcagcaa	5341
	ggttgatggg	ctcgagttag	tattgaaggt	ggggggatga	atgggggttg	gatggggcta	5401
	tggctgggaa	gggggcttcg	gtgggagacg	tgggaagaggt	tgggaagcaga	gcgatgtttc	5461
10	ttcatcctca	aaggtgtcac	tcacctctcc	cacctatgtc	tcccccgacc	tttctctcct	5521
	caactactgt	ctctcccacc	tcagccgcta	catagttcac	ctggggcagc	acaacctcca	5581
	gaaggaggag	ggctgtgagc	agaccoggac	agccactgag	tccttcccc	accccggtt	5641
	caacaacagc	ctcccccaaca	aagaccaccg	caatgacatc	atgctggtga	agatggcatc	5701
	gccagtctcc	atcacctggg	ctgtgcgacc	cctcacctc	tcctcacgct	gtgtcactgc	5761
15	tggcaccagc	tgctcatatt	ccggctgggg	cagcacgtcc	agccccagt	gtaggagcac	5821
	cagaggggaa	cctggcaggg	ggtggtgagg	agggagtgg	caggattgtg	gaaggggtca	5881
	gggcatcaga	gatgcggttc	acagtgcaga	tgtgggataa	gttgagagga	tgtgtggaaa	5941
	acgtcaggat	aggggggtgg	ggacaaaagt	tggggccttg	gagtcagacg	gacgggatat	6001
	gcaatcatac	atccataacc	tcctggttgt	aagaccttag	gcaagcagct	tcacctctct	6061
20	gaatcttgat	tttcttctct	ataaaatgag	aatgattata	cccacctgtc	aggattggat	6121
	tagagataat	gtatatcaag	caactgacat	aaatcattta	ttggatagca	ggctgggcac	6181
	cgtggctcac	gcctgtaatc	ccagcacttt	gggaggccga	ggtgggaaga	tcacctgagg	6241
	tcaggacttt	gataccagcc	tggccaacgt	ggtgaaatcc	catctctact	aaaaatgtga	6301
	aaattagtgt	ggcgtgggtg	tgtgcgcctg	taatcccagc	tactcgggag	gttgaggcag	6361
25	gagaatcgct	taaacttggg	agacggaggt	tgacgtgagc	caagatcacg	ccactgcact	6421
	ccagcctggg	caacagagca	agactctgtc	tcgaaaaaaa	aaaaaaaaaa	gctggatagc	6481
	attgctgttg	ctattgttac	aagaagagag	gtgagttggc	tgctgtctaa	gacagggatt	6541
	ccccagggg	cgggatcaca	gcaagcactg	cattagggga	ggtggcaggg	ggctcattcc	6601
	cacagccctt	cacgctgttt	ccacagtacg	cctgcctcac	accttgcgat	gcgccaacat	6661
30	caccatcatt	gagcaccaga	agtgtgagaa	cgcctacccc	ggcaacatca	cagacaccat	6721
	ggtgtgtgcc	agcgtgcagg	aagggggcaa	ggactcctgc	caggctcagt	tgggtctcaa	6781
	ccacagcccc	atccccatcc	ccagcttcaa	tgacatcttt	accgacatcc	acaatttcat	6841
	ccccaacctc	aacccgcccga	ccoctgcaac	tcccaatcca	tctcttcccc	tgttcccggt	6901
	tctgacctca	gcacaaaactt	cagctccatc	cccgttttcca	caccatttcc	agctccaacc	6961
35	atccccaaac	tcgtttttga	gcctaaccct	atcctttatc	ccaaccataa	tcccagcttt	7021
	atcgctaaac	ctatcacctt	tcccagtgcc	taccatcct	gtctcggccc	cactctctaa	7081
	caccgtcccc	acctcctccc	tggctaaccac	catgctcaac	gctttctctg	accgacattc	7141
	tctctccccg	tgcccagggt	gactccgggg	gccctctggt	ctgtaaccag	tctcttcaag	7201
	gcattatctc	ctggggccag	gatccgtgtg	cgatcacccg	aaagcctggt	gtctacacga	7261
40	aagtctgcaa	atatgtggac	tggatccagg	agacgatgaa	gaacaattag	actggaccca	7321
	cccaccacag	cccatcacc	tccatttcca	cttgggtgtt	ggttcctggt	cactctgtta	7381
	ataagaaaac	ctaagccaag	acctctacg	aacattcttt	gggcctcctg	gactacagga	7441
	gatgctgtca	cttaataatc	aacctggggt	tcgaaatcag	tgagacctgg	attcaaatcc	7501
	tgcttgaaa	tattgtgact	ctgggaatga	caacacctgg	tttgttctct	gttgtatccc	7561
45	cagccccaaa	gacagctcct	ggccatatat	caaggtttca	ataaatattt	gctaaatgag	7621
	tgaatctact	gagtgtttac	tatgtgctag	accctgatcc	aatggctttt	attttatattt	7681
	attttttgag	agagtctcgc	tctgtcacc	aggctggagt	acagtgggtc	tatctctgct	7741
	cactgcaacc	tccacctcct	gggttcaagc	aattctcctg	cctcagcctc	ctgaatagct	7801
	gggattacag	gtgcctacca	ccacatccgg	ctaatttttg	tatttttttag	tagagatggg	7861
50	gcttcacat	gttggccagg	ctggtctcga	actcctgacc	tcagatgatc	tgccctcctt	7921
	ggcctcccaa	actcctggga	ttacagacgt	gagccaccgc	gcccggccgg	ctttcattta	7981
	ttaattaaaa	gaaattaaat	taattaatct	athtaggaga	cagtcttgct	ctggtgcccc	8041
	ggctggagtg	cagtaacaat	cacagctcac	agcaatctca	atttcctggg	gtcaagtgtat	8101
	tgctctcct	cagcctccag	agtagctggg	actaacggca	catgccacga	agcccagcta	8161
55	atttttgtat	ttttcgtaga	gacagaggtc	tcagtatgtt	gccccggcta	gtctcaaaact	8221
	cctgggctca	agcagtctgt	cctcctcagc	ctccaaaagt	ggtgagatta	caggcatgag	8281
	tcgctgtgcc	tggcctccaa	gcactttcaa	atgtatcaac	ttaatcctca	caaaaccttg	8341
	tgaggtcggt	actgttttca	tacctatttt	atagttgaag	aaacagacac	agagaagcaa	8401
	agtcacttgc	tcacagtcac	gtggctagga	gagcaaggat	ctgaagcaag	gcgatctctt	8461

```

aattaccaag tgatgttcct ggagtaaggc tctgtttggt tccttttcctg taaaatgctg 8521
catgcaaaag tataacacag taagtaaaga agtcagttag cctgcacata ctaagacctg 8581
accaaaggag cttattgttt ttctccaact tccatgatag gtaattagat agtggagacc 8641
tctgctggcc aatatggtag ccactaaccg cagctggctc ttccaattaa aattacataa 8701
5 agccagaaat gtaactcctc tgtctcactt gttatatctc caaggctgga tagccacatg 8761
tgactgggtg tggctggatt agctagtgcg tataaaacat cactgcagaa agttcagctg 8821
agcagcactg agttagatgg cctctgaaga ggatgtccca cggagagaat ccagaactca 8881
ggatcctttt tttttttttt ctttgcgaca gagtcttgct ctgtcaccga ggctggagtg 8941
cagtggcgtg atctcggctc actgcaactt ctgcctccca ggttcaagca attctcctgc 9001
10 ctcagcctcc ctagtagctg ggactacagg cctgtgcca catccccagc taatttttgt 9061
gtcttttttag tagagatggg gtttcactat gttggccagg ctggtctcga actcctgacc
```